

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or priority of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Type of Search

Vendors and cost where applicable

Searcher BA Beverly C4994 NA Sequence (#) _____ STN _____
 Searcher Phone # _____ AA Sequence (#) _____ Dialog _____
 Searcher Location _____ Structure (#) _____ Questel/Orbit _____
 Date Completed 02-25-00 Bibliographic _____ Dr Link _____
 Searcher Prep & Review Time _____ Litigation _____ Lexis Nexis _____
 Clerical Prep Time _____ Patent Family _____ Sequence Systems _____
 Online Time _____ Other _____ WWW Internet _____
 Other (specify) MP



[illegible]

674	31	100	0	892	US-60-174	Sequence	4184	Applic	980-03
680	31	100	0	893	US-60-174	Sequence	4184	Applic	980-03
681	31	100	0	894	US-60-174	Sequence	4184	Applic	980-03
682	31	100	0	895	US-60-174	Sequence	4184	Applic	980-03
683	31	100	0	896	US-60-174	Sequence	4184	Applic	980-03
684	31	100	0	897	US-60-174	Sequence	4184	Applic	980-03
685	31	100	0	898	US-60-174	Sequence	4184	Applic	980-03
686	31	100	0	899	US-60-174	Sequence	4184	Applic	980-03
687	31	100	0	900	US-60-174	Sequence	4184	Applic	980-03
688	31	100	0	901	US-60-174	Sequence	4184	Applic	980-03
689	31	100	0	902	US-60-174	Sequence	4184	Applic	980-03
690	31	100	0	903	US-60-174	Sequence	4184	Applic	980-03
691	31	100	0	904	US-60-174	Sequence	4184	Applic	980-03
692	31	100	0	905	US-60-174	Sequence	4184	Applic	980-03
693	31	100	0	906	US-60-174	Sequence	4184	Applic	980-03
694	31	100	0	907	US-60-174	Sequence	4184	Applic	980-03
695	31	100	0	908	US-60-174	Sequence	4184	Applic	980-03
696	31	100	0	909	US-60-174	Sequence	4184	Applic	980-03
697	31	100	0	910	US-60-174	Sequence	4184	Applic	980-03
698	31	100	0	911	US-60-174	Sequence	4184	Applic	980-03
699	31	100	0	912	US-60-174	Sequence	4184	Applic	980-03
700	31	100	0	913	US-60-174	Sequence	4184	Applic	980-03
701	31	100	0	914	US-60-174	Sequence	4184	Applic	980-03
702	31	100	0	915	US-60-174	Sequence	4184	Applic	980-03
703	31	100	0	916	US-60-174	Sequence	4184	Applic	980-03
704	31	100	0	917	US-60-174	Sequence	4184	Applic	980-03
705	31	100	0	918	US-60-174	Sequence	4184	Applic	980-03
706	31	100	0	919	US-60-174	Sequence	4184	Applic	980-03
707	31	100	0	920	US-60-174	Sequence	4184	Applic	980-03
708	31	100	0	921	US-60-174	Sequence	4184	Applic	980-03
709	31	100	0	922	US-60-174	Sequence	4184	Applic	980-03
710	31	100	0	923	US-60-174	Sequence	4184	Applic	980-03
711	31	100	0	924	US-60-174	Sequence	4184	Applic	980-03
712	31	100	0	925	US-60-174	Sequence	4184	Applic	980-03
713	31	100	0	926	US-60-174	Sequence	4184	Applic	980-03
714	31	100	0	927	US-60-174	Sequence	4184	Applic	980-03
715	31	100	0	928	US-60-174	Sequence	4184	Applic	980-03
716	31	100	0	929	US-60-174	Sequence	4184	Applic	980-03
717	31	100	0	930	US-60-174	Sequence	4184	Applic	980-03
718	31	100	0	931	US-60-174	Sequence	4184	Applic	980-03
719	31	100	0	932	US-60-174	Sequence	4184	Applic	980-03
720	31	100	0	933	US-60-174	Sequence	4184	Applic	980-03
721	31	100	0	934	US-60-174	Sequence	4184	Applic	980-03
722	31	100	0	935	US-60-174	Sequence	4184	Applic	980-03
723	31	100	0	936	US-60-174	Sequence	4184	Applic	980-03
724	31	100	0	937	US-60-174	Sequence	4184	Applic	980-03
725	31	100	0	938	US-60-174	Sequence	4184	Applic	980-03
726	31	100	0	939	US-60-174	Sequence	4184	Applic	980-03
727	31	100	0	940	US-60-174	Sequence	4184	Applic	980-03
728	31	100	0	941	US-60-174	Sequence	4184	Applic	980-03
729	31	100	0	942	US-60-174	Sequence	4184	Applic	980-03
730	31	100	0	943	US-60-174	Sequence	4184	Applic	980-03
731	31	100	0	944	US-60-174	Sequence	4184	Applic	980-03
732	31	100	0	945	US-60-174	Sequence	4184	Applic	980-03
733	31	100	0	946	US-60-174	Sequence	4184	Applic	980-03
734	31	100	0	947	US-60-174	Sequence	4184	Applic	980-03
735	31	100	0	948	US-60-174	Sequence	4184	Applic	980-03
736	31	100	0	949	US-60-174	Sequence	4184	Applic	980-03
737	31	100	0	950	US-60-174	Sequence	4184	Applic	980-03
738	31	100	0	951	US-60-174	Sequence	4184	Applic	980-03
739	31	100	0	952	US-60-174	Sequence	4184	Applic	980-03
740	31	100	0	953	US-60-174	Sequence	4184	Applic	980-03
741	31	100	0	954	US-60-174	Sequence	4184	Applic	980-03
742	31	100	0	955	US-60-174	Sequence	4184	Applic	980-03
743	31	100	0	956	US-60-174	Sequence	4184	Applic	980-03
744	31	100	0	957	US-60-174	Sequence	4184	Applic	980-03
745	31	100	0	958	US-60-174	Sequence	4184	Applic	980-03
746	31	100	0	959	US-60-174	Sequence	4184	Applic	980-03
747	31	100	0	960	US-60-174	Sequence	4184	Applic	980-03
748	31	100	0	961	US-60-174	Sequence	4184	Applic	980-03
749	31	100	0	962	US-60-174	Sequence	4184	Applic	980-03
750	31	100	0	963	US-60-174	Sequence	4184	Applic	980-03


```

971 24 92 5 720 3 US-60-150- Sequence 482, Applicat 5,84e+03
972 24 92 5 721 3 US-60-151- Sequence 1521, Applic 5,84e+03
973 24 92 5 722 3 US-60-152- Sequence 4023, Applic 5,84e+03
974 24 92 5 723 3 US-60-153- Sequence 1, Applic 5,84e+03
975 24 92 5 724 3 US-60-154- Sequence 8421, Applic 5,84e+03
976 24 92 5 725 3 US-60-155- Sequence 2622, Applic 5,84e+03
977 24 92 5 726 3 US-60-156- Sequence 15113, Applic 5,84e+03
978 24 92 5 727 3 US-60-157- Sequence 20513, Applic 5,84e+03
979 24 92 5 728 3 US-60-158- Sequence 0018, Applic 5,84e+03
980 24 92 5 729 3 US-60-159- Sequence 11231, Applic 5,84e+03
981 24 92 5 730 3 US-60-160- Sequence 1522, Applic 5,84e+03
982 24 92 5 731 3 US-60-161- Sequence 59116, Applic 5,84e+03
983 24 92 5 732 3 US-60-162- Sequence 6191, Applic 5,84e+03
984 24 92 5 733 3 US-60-163- Sequence 2010, Applic 5,84e+03
985 24 92 5 734 3 US-60-164- Sequence 202, Applicat 5,84e+03
986 24 92 5 735 3 US-60-165- Sequence 42434, Applic 5,84e+03
987 24 92 5 736 3 US-60-166- Sequence 18229, Applic 5,84e+03
988 24 92 5 737 3 US-60-167- Sequence 3863, Applic 5,84e+03
989 24 92 5 738 3 US-60-168- Sequence 1253, Applic 5,84e+03
990 24 92 5 739 3 US-60-169- Sequence 14157, Applic 5,84e+03
991 24 92 5 740 3 US-60-170- Sequence 14379, Applic 5,84e+03
992 24 92 5 741 3 US-60-171- Sequence 10045, Applic 5,84e+03
993 24 92 5 742 3 US-60-172- Sequence 384, Applicat 5,84e+03
994 24 92 5 743 3 US-60-173- Sequence 170, Applicat 5,84e+03
995 24 92 5 744 3 US-60-174- Sequence 5189, Applic 5,84e+03
996 24 92 5 745 3 US-60-175- Sequence 14533, Applic 5,84e+03
997 24 92 5 746 3 US-60-176- Sequence 2250, Applic 5,84e+03
998 24 92 5 747 3 US-60-177- Sequence 17102, Applic 5,84e+03
999 24 92 5 748 3 US-60-178- Sequence 17102, Applic 5,84e+03
1000 24 92 5 749 3 US-60-179- Sequence 17102, Applic 5,84e+03

```

ALIGNMENTS

```

RESULT 1
ID US-09-132-799-1 STANFAPD: PPT: 4 AA.
XX AC xxxxxx
XX DT
XX DI
XX DE
XX SE Sequence 1, Application US/09132799
XX SE Sequence 1, Application US/09132799
CC SE Sequence 1, Application US/09132799
CC GENRAL INFORMATION:
CC APPLICANT: Schenrock, Uwe
CC APPLICANT: Max, Heiner
CC APPLICANT: Hearing, Vincent J
CC TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL PREPARATIONS COMPRISING
CC TITLE OF INVENTION: OLIGOPEPTIDES FOR LIGHTENING THE SKIN OF AGE MARKS
CC TITLE OF INVENTION: AND/OR FOR PREVENTING TANNING OF THE SKIN, IN
CC TITLE OF INVENTION: PARTICULAR TANNING OF THE SKIN CAUSED BY UV RADIATIO
CC FILE REFERENCE: Patsdorff-514-KGB
CC CURRENT APPLICATION NUMBER: US/09/132-799
CC CURRENT FILING DATE: 1998-08-13
CC NUMBER OF SEQ ID NOS: 3
CC SOFTWARE: Patent In Ver. 2.1
CC SEQ ID NO 1
CC LENGTH: 4
CC TYPE: PPT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Structural
CC OTHER INFORMATION: Base for oligopeptides
CC OTHER INFORMATION: base for oligopeptides.
SQ SEQUENCE 4 AA: 470 MW: 126.00
Query Match 100.0% Score 31 DB 16 Length 4
Best Local Similarity 100.0% Prod. No. 2.98e-03
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 1 VWRP 4
1111
Cy 1 VWRP 4

```

```

RESULT 2
ID US-09-132-799-2 STANFAPD: PPT: 5 AA.
XX AC xxxxxx
XX DT
XX DI
XX DE
XX SE Sequence 2, Application US/09132799
XX SE Sequence 2, Application US/09132799
CC SE Sequence 2, Application US/09132799
CC GENRAL INFORMATION:
CC APPLICANT: Schenrock, Uwe
CC APPLICANT: Max, Heiner
CC APPLICANT: Hearing, Vincent J
CC TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL PREPARATIONS COMPRISING
CC TITLE OF INVENTION: OLIGOPEPTIDES FOR LIGHTENING THE SKIN OF AGE MARKS
CC TITLE OF INVENTION: AND/OR FOR PREVENTING TANNING OF THE SKIN, IN
CC TITLE OF INVENTION: PARTICULAR TANNING OF THE SKIN CAUSED BY UV RADIATIO
CC FILE REFERENCE: Patsdorff-514-KGB
CC CURRENT APPLICATION NUMBER: US/09/132-799
CC CURRENT FILING DATE: 1998-08-13
CC NUMBER OF SEQ ID NOS: 3
CC SOFTWARE: Patent In Ver. 2.1
CC SEQ ID NO 2
CC LENGTH: 5
CC TYPE: PPT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Structural base
CC OTHER INFORMATION: for oligopeptides
SQ SEQUENCE 5 AA: 567 MW: 201.00
Query Match 100.0% Score 31 DB 16 Length 5
Best Local Similarity 100.0% Prod. No. 2.98e-03
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 1 VWRP 4
1111
Cy 1 VWRP 4

```

```

RESULT 3
ID US-09-132-799-3 STANFAPD: PPT: 6 AA.
XX AC xxxxxx
XX DT
XX DI
XX DE
XX SE Sequence 3, Application US/09132799
XX SE Sequence 3, Application US/09132799
CC SE Sequence 3, Application US/09132799
CC GENRAL INFORMATION:
CC APPLICANT: Schenrock, Uwe
CC APPLICANT: Max, Heiner
CC APPLICANT: Hearing, Vincent J
CC TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL PREPARATIONS COMPRISING
CC TITLE OF INVENTION: OLIGOPEPTIDES FOR LIGHTENING THE SKIN OF AGE MARKS
CC TITLE OF INVENTION: AND/OR FOR PREVENTING TANNING OF THE SKIN, IN
CC TITLE OF INVENTION: PARTICULAR TANNING OF THE SKIN CAUSED BY UV RADIATIO
CC FILE REFERENCE: Patsdorff-514-KGB
CC CURRENT APPLICATION NUMBER: US/09/132-799
CC CURRENT FILING DATE: 1998-08-13
CC NUMBER OF SEQ ID NOS: 3
CC SOFTWARE: Patent In Ver. 2.1
CC SEQ ID NO 3
CC LENGTH: 6
CC TYPE: PPT
CC ORGANISM: Artificial Sequence
CC FEATURE:
CC OTHER INFORMATION: Description of Artificial Sequence: Structural
CC OTHER INFORMATION: base for oligopeptides.
SQ SEQUENCE 6 AA: 567 MW: 201.00
Query Match 100.0% Score 31 DB 16 Length 5
Best Local Similarity 100.0% Prod. No. 2.98e-03
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 1 VWRP 4
1111
Cy 1 VWRP 4

```

XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	DE
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	GENERAL INFORMATION:
CC	APPLICANT: Solaris, Peter J.
CC	APPLICANT: Chell, William G.
CC	APPLICANT: Miller, Jeff F.
CC	APPLICANT: Stenger, William P.C.
CC	TITLE OF INVENTION: Peptide Library and Screening Method
CC	NUMBER OF SEQUENCES: 71
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: William M. Smith
CC	STREET: One Market Plaza, Stewart Tower, Suite 2000
CC	CITY: San Francisco
CC	STATE: California
CC	COUNTRY: USA
CC	ZIP: 94105
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPILER: IBM PC compatible
CC	OPERATING SYSTEM: DOS/MS-DOS
CC	SOFTWARE: Patent Release #1.0, Version #1.0
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/77892/0879
CC	FILING DATE: 1992/06/16
CC	CLASSIFICATION:
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US/7778/223
CC	FILING DATE: 16-OCT-1991
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Smith, William M.
CC	REGISTRATION NUMBER: 40,223
CC	REFERENCE/CITATION NUMBER: 1150q-50-a
CC	TELEPHONE: 415-426-2400
CC	TELEFAX: 415-426-2422
CC	INFORMATION FOR SEQ ID NO: 32:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 12 amino acids
CC	TYPE: AMINO ACIDS
CC	STRANDEDNESS: Single
CC	TOPOLOGY: Linear
CC	MOLECULE TYPE: peptide
CC	IMMEDIATE SOURCE:
CC	CLONE: C9.2.1.1
CC	SEQUENCE 12 AA: 1576 MW: 912 GN:
XX	Query Match 100.0% Score 31 DB 11 Length 12
XX	Best Local Similarity 100.0% Prod. No. 2,980-031
XX	Matches 4 Conserved 0 Mismatches 0
D6	8 VWRP 11
OZ	1 VWRP 4
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	DE
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	GENERAL INFORMATION:
CC	APPLICANT: Solaris, Peter J.
CC	APPLICANT: Chell, William G.
CC	APPLICANT: Miller, Jeff F.
CC	APPLICANT: Stenger, William P.C.
CC	TITLE OF INVENTION: Peptide Library and Screening Method
CC	NUMBER OF SEQUENCES: 71
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: William M. Smith
CC	STREET: One Market Plaza, Stewart Tower, Suite 2000
CC	CITY: San Francisco
CC	STATE: California
CC	COUNTRY: USA
CC	ZIP: 94105
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPILER: IBM PC compatible
CC	OPERATING SYSTEM: DOS/MS-DOS
CC	SOFTWARE: Patent Release #1.0, Version #1.0
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/77892/0879
CC	FILING DATE: 16-OCT-1991
CC	CLASSIFICATION:
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US/7778/223
CC	FILING DATE: 16-OCT-1991
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Smith, William M.
CC	REGISTRATION NUMBER: 40,223
CC	REFERENCE/CITATION NUMBER: 1150q-50-a
CC	TELEPHONE: 415-426-2400
CC	TELEFAX: 415-426-2422
CC	INFORMATION FOR SEQ ID NO: 32:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 12 amino acids
CC	TYPE: AMINO ACIDS
CC	STRANDEDNESS: Single
CC	TOPOLOGY: Linear
CC	MOLECULE TYPE: peptide
CC	IMMEDIATE SOURCE:
CC	CLONE: C9.2.1.1
CC	SEQUENCE 12 AA: 1576 MW: 912 GN:
XX	Query Match 100.0% Score 31 DB 5 Length 12
XX	Best Local Similarity 100.0% Prod. No. 2,980-031
XX	Matches 4 Conserved 0 Mismatches 0
D6	8 VWRP 11
OZ	1 VWRP 4
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	DE
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	GENERAL INFORMATION:
CC	APPLICANT: Solaris, Peter J.
CC	APPLICANT: Chell, William G.
CC	APPLICANT: Miller, Jeff F.
CC	APPLICANT: Stenger, William P.C.
CC	TITLE OF INVENTION: Peptide Library and Screening Method
CC	NUMBER OF SEQUENCES: 71
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: William M. Smith
CC	STREET: One Market Plaza, Stewart Tower, Suite 2000
CC	CITY: San Francisco
CC	STATE: California
CC	COUNTRY: USA
CC	ZIP: 94105
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPILER: IBM PC compatible
CC	OPERATING SYSTEM: DOS/MS-DOS
CC	SOFTWARE: Patent Release #1.0, Version #1.0
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/77892/0879
CC	FILING DATE: 16-OCT-1991
CC	CLASSIFICATION:
CC	PRIOR APPLICATION DATA:
CC	APPLICATION NUMBER: US/7778/223
CC	FILING DATE: 16-OCT-1991
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: Smith, William M.
CC	REGISTRATION NUMBER: 40,223
CC	REFERENCE/CITATION NUMBER: 1150q-50-a
CC	TELEPHONE: 415-426-2400
CC	TELEFAX: 415-426-2422
CC	INFORMATION FOR SEQ ID NO: 32:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 12 amino acids
CC	TYPE: AMINO ACIDS
CC	STRANDEDNESS: Single
CC	TOPOLOGY: Linear
CC	MOLECULE TYPE: peptide
CC	IMMEDIATE SOURCE:
CC	CLONE: C9.2.1.1
CC	SEQUENCE 12 AA: 1576 MW: 912 GN:
XX	Query Match 100.0% Score 31 DB 5 Length 12
XX	Best Local Similarity 100.0% Prod. No. 2,980-031
XX	Matches 4 Conserved 0 Mismatches 0
D6	8 VWRP 11
OZ	1 VWRP 4
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	DE
XX	SEQUENCE 32, APPLICATION US/77892/0879
XX	GENERAL INFORMATION:
CC	APPLICANT: Solaris, Peter J.
CC	APPLICANT: Chell, William G.
CC	APPLICANT: Miller, Jeff F.
CC	APPLICANT: Stenger, William P.C.
CC	TITLE OF INVENTION: Peptide Library and Screening Method
CC	NUMBER OF SEQUENCES: 71
CC	CORRESPONDENCE ADDRESS:

CC GENERAL INFORMATION:
 CC APPLICANT: Schatz, Peter C.
 CC APPLICANT: Cull, Millard G.
 CC APPLICANT: Miller, Jeff F.
 CC APPLICANT: Stemmer, William P.C.
 CC APPLICANT: Gates, Christian M.
 CC TITLE OF INVENTION: Peptide Library and Screening Method
 CC NUMBER OF SEQUENCES: 152
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: William M. Smith
 CC STREET: One Market Plaza, Steuart Tower, Suite 2000
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 94104
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.20
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09/010,216
 CC FILING DATE:
 CC CLASSIFICATION:
 CC PRACTICE APPLICATION DATA:
 CC APPLICATION NUMBER: US 08,549,341
 CC FILING DATE: 26-08-1995
 CC APPLICATION NUMBER: US 08,290,641
 CC FILING DATE: 15 AUG 1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/963,141
 CC FILING DATE: 15-OCT-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Smith, William M.
 CC REGISTRATION NUMBER: 30,223
 CC REFERENCE/DOCKET NUMBER: 16,487,014-105
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415-926-2400
 CC TELEFAX: 415-926-2422
 CC INFORMATION FOR SEQ ID NO. 32
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 12 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC IMMEDIATE SOURCE:
 CC CLONE: 59-2-1-1
 CC SEQUENCE 12 AA: 1376 MW: 912 CN:
 CC
 CC Query Match 100.0% Score 31: DB 15: Length 12:
 CC Best Local Similarity 100.0% Pctd No. 2 99C-03:
 CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 8 VWRP 11
 Q1 1 VWRP 4

RESULT 7
 ID US-09-622-632-15 STANDARD: PRI: 15 AA
 XX AC xxxxxx
 XX DI
 XX SEQUENCE 30: Application US/09/010,216
 XX
 XX Sequence 32: Application US/09/010,216
 XX GENERAL INFORMATION:
 XX APPLICANT: Anthony C. Forster
 XX TITLE OF INVENTION: Immunologic Screening Process, and
 XX TITLE OF INVENTION: Compositions Thereof

CC NUMBER OF SEQUENCES: 106
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: LAHIVE & COCKFIELD
 CC STREET: 60 State Street
 CC CITY: Boston
 CC STATE: Massachusetts
 CC COUNTRY: USA
 CC ZIP: 02109-1975
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09/010,216
 CC FILING DATE: 21-MARCH-1996
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER:
 CC FILING DATE:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Vincent, Matthew P.
 CC REGISTRATION NUMBER: 36,104
 CC REFERENCE/DOCKET NUMBER: HMT-018
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 1647-227-7400
 CC TELEFAX: 1647-227-5941
 CC INFORMATION FOR SEQ ID NO. 33:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 15 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: linear
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC IMMEDIATE SOURCE:
 CC CLONE: 59-2-1-1
 CC SEQUENCE 15 AA: 1602 MW: 1180 CN:
 CC
 CC Query Match 100.0% Score 31: DB 15: Length 15:
 CC Best Local Similarity 100.0% Pctd No. 2 99C-03:
 CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 8 VWRP 11
 Q1 1 VWRP 4

RESULT 8
 ID US-08-488-315A-8 STANDARD: PRI: 16 AA
 XX AC xxxxxx
 XX DI
 XX SEQUENCE 5: Application US/08/488,315A
 XX
 XX Sequence 8: Application US/08/488,315A
 XX GENERAL INFORMATION:
 XX APPLICANT: Atassi, M. Zouhair
 XX TITLE OF INVENTION: Detection and Treatment of Myasthenia
 XX TITLE OF INVENTION: Graves
 XX NUMBER OF SEQUENCES: 25
 XX CORRESPONDENCE ADDRESS:
 XX ADDRESSEE: C. Steven McDaniel
 XX STREET: P.O. Box 3267
 XX CITY: Houston
 XX STATE: Texas
 XX COUNTRY: USA
 XX ZIP: 77052-2267
 XX COMPUTER READABLE FORM:
 XX MEDIUM TYPE: Floppy disk
 XX COMPUTER: IBM PC compatible
 XX OPERATING SYSTEM: MS-DOS
 XX SOFTWARE: WordPerfect 5.1
 XX CURRENT APPLICATION DATA:
 XX APPLICATION NUMBER: US/08/488,315A

CC FILING DATE: 17 JUNE 1995
 CC CLASSIFICATION: 614
 CC APPLICATION INFORMATION:
 CC NAME: MCGILL, C. Steven
 CC REFERENCE NUMBER: 42,714
 CC REFERENCE NUMBER: 1516-00408
 CC TELEPHONE: 713/248-8000
 CC TELEFAX: 713/248-8008
 CC INFORMATION FOR SEQ ID NO: 8:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 16 amino acids
 CC TYPE: amino acid
 CC PRT: 118004
 CC M OTYPE TYPE: poplido
 CC SEQUENCE: 16 AA: 1902 MW: 1476 CN:
 Query Match 100.0% Score 31: DB 9: Length 16:
 Best Local Similarity 100.0% Pred. No. 2,980-03:
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 7 VWRP 16
 QY 1 VWRP 4
 RESULT 10
 ID US-99-170-440-1-094 STANDARD: PRT: 1, AA:
 XX
 AC
 XX
 AC
 XX
 DT
 XX
 DE Sequence 1389, Application US/60170346
 XX
 CC Sequence 1389, Application US/60170346
 CC GENERAL INFORMATION:
 CC APPLICANT: BODAZZ, Vivian
 CC TITLE OF INVENTION: ISOLATED HUMAN GPCR PROTEIN, NOVELLY
 CC TITLE OF INVENTION: ADRI M. BODAZZ, INVENTOR; BODAZZ, ADRI M., ATTORNEY
 CC FILE REFERENCE: 121-440143
 CC CURRENT FILING DATE: 1999-12-14
 CC NUMBER OF SEQ ID NOS: 2600
 CC SOFTWARE: FastSeq for Windows Version 4.0
 CC SEQ ID NO: 1389
 CC LENGTH: 31
 CC TYPE: PRT
 CC ORGANISM: HUMAN
 CC SEQUENCE: 31 AA: 1417 MW: 4529 CN:
 Query Match 100.0% Score 31: DB 26: Length 31:
 Best Local Similarity 100.0% Pred. No. 2,980-04:
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 25 VWRP 28
 QY 1 VWRP 4
 RESULT 11
 ID PCI-US99-26404-92 STANDARD: PRT: 14 AA:
 XX
 AC
 XX
 AC
 XX
 DT
 XX
 DE Sequence 92, Application PCT/US99/26404
 XX
 CC Sequence 92, Application PCT/US99/26404
 CC GENERAL INFORMATION:
 CC APPLICANT: human genome sciences, Inc., et al
 CC TITLE OF INVENTION: A Human Secreted Protein
 CC FILE REFERENCE: 121-440143
 CC CURRENT FILING DATE: 1999-11-04
 CC EARLIER FILING DATE: 1999-11-04
 CC EARLIER FILING DATE: 1999-11-04
 CC NUMBER OF SEQ ID NOS: 115
 CC SOFTWARE: Patent In Ver. 2.0
 CC SEQ ID NO: 92
 CC LENGTH: 34
 CC TYPE: PRT
 CC ORGANISM: Homo sapiens
 CC SEQUENCE: 34 AA: 2763 MW: 3174 CN:
 Query Match 100.0% Score 31: DB 11: Length 34:
 Best Local Similarity 100.0% Pred. No. 2,980-05:
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 14 VWRP 17

CC FILING DATE: 17 JUNE 1995
 CC CLASSIFICATION: 614
 CC APPLICATION INFORMATION:
 CC NAME: MCGILL, C. Steven
 CC REFERENCE NUMBER: 42,714
 CC REFERENCE NUMBER: 1516-00408
 CC TELEPHONE: 713/248-8000
 CC TELEFAX: 713/248-8008
 CC INFORMATION FOR SEQ ID NO: 8:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 16 amino acids
 CC TYPE: amino acid
 CC PRT: 118004
 CC M OTYPE TYPE: poplido
 CC SEQUENCE: 16 AA: 1902 MW: 1476 CN:
 Query Match 100.0% Score 31: DB 9: Length 16:
 Best Local Similarity 100.0% Pred. No. 2,980-03:
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 7 VWRP 16
 QY 1 VWRP 4
 RESULT 10
 ID US-99-170-440-1-094 STANDARD: PRT: 1, AA:
 XX
 AC
 XX
 AC
 XX
 DT
 XX
 DE Sequence 1389, Application US/60170346
 XX
 CC Sequence 1389, Application US/60170346
 CC GENERAL INFORMATION:
 CC APPLICANT: BODAZZ, Vivian
 CC TITLE OF INVENTION: ISOLATED HUMAN GPCR PROTEIN, NOVELLY
 CC TITLE OF INVENTION: ADRI M. BODAZZ, INVENTOR; BODAZZ, ADRI M., ATTORNEY
 CC FILE REFERENCE: 121-440143
 CC CURRENT FILING DATE: 1999-12-14
 CC NUMBER OF SEQ ID NOS: 2600
 CC SOFTWARE: FastSeq for Windows Version 4.0
 CC SEQ ID NO: 1389
 CC LENGTH: 31
 CC TYPE: PRT
 CC ORGANISM: HUMAN
 CC SEQUENCE: 31 AA: 1417 MW: 4529 CN:
 Query Match 100.0% Score 31: DB 26: Length 31:
 Best Local Similarity 100.0% Pred. No. 2,980-04:
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 25 VWRP 28
 QY 1 VWRP 4
 RESULT 11
 ID PCI-US99-26404-92 STANDARD: PRT: 14 AA:
 XX
 AC
 XX
 AC
 XX
 DT
 XX
 DE Sequence 92, Application PCT/US99/26404
 XX
 CC Sequence 92, Application PCT/US99/26404
 CC GENERAL INFORMATION:
 CC APPLICANT: human genome sciences, Inc., et al
 CC TITLE OF INVENTION: A Human Secreted Protein
 CC FILE REFERENCE: 121-440143
 CC CURRENT FILING DATE: 1999-11-04
 CC EARLIER FILING DATE: 1999-11-04
 CC EARLIER FILING DATE: 1999-11-04
 CC NUMBER OF SEQ ID NOS: 115
 CC SOFTWARE: Patent In Ver. 2.0
 CC SEQ ID NO: 92
 CC LENGTH: 34
 CC TYPE: PRT
 CC ORGANISM: Homo sapiens
 CC SEQUENCE: 34 AA: 2763 MW: 3174 CN:
 Query Match 100.0% Score 31: DB 11: Length 34:
 Best Local Similarity 100.0% Pred. No. 2,980-05:
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 14 VWRP 17

DB 12 VWRP 15
 QY 1 VWRP 4

RESULT 12
 ID US-09-150-202-3702 STANDARD: PRI: 40 AA.
 XX XXXXXX

Sequence 12, Application PC/US99030309
 GENERAL INFORMATION:
 APPLICANT: Human Genome Sciences, Inc.
 TITLE OF INVENTION: 36 Human Secreted Proteins
 FILE REFERENCE: C1000114
 CURRENT FILING DATE: 1998-02-24
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,051
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,053
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,054
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,052
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,051
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,050
 EARLIER FILING DATE: 1998-02-24
 NUMBER OF SEQ ID NOS: 176
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO: 35
 LENGTH: 35
 TYPE: PRT
 ORGANISM: HOMO SAPIENS

Query Match: 100.0% Score 31 DB 1 Length 35.
 Best Local Similarity 100.0% Prod No. 2598-03
 Matches 4 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 12 VWRP 15
 QY 1 VWRP 4

RESULT 13
 ID US-09-150-202-3702 STANDARD: PRI: 40 AA.
 XX XXXXXX

Sequence 12, Application PC/US99030309
 GENERAL INFORMATION:
 APPLICANT: HUMAN GENOME SCIENCES, INC.
 TITLE OF INVENTION: 36 HUMAN SECRETED PROTEINS
 FILE REFERENCE: C1000114
 CURRENT FILING DATE: 1998-02-24
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,051
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,053
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,054
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,052
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,051
 EARLIER FILING DATE: 1998-02-24
 EARLIER APPLICATION NUMBER: 69/076,050
 EARLIER FILING DATE: 1998-02-24
 NUMBER OF SEQ ID NOS: 176
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO: 35
 LENGTH: 35
 TYPE: PRT
 ORGANISM: HUMAN

Query Match: 100.0% Score 31 DB 2 Length 40.
 Best Local Similarity 100.0% Prod No. 2598-03
 Matches 4 Conservative 0 Mismatches 0 Indels 0 Gaps 0

Matches 4: Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 24 VWRP 27
 QY 1 VWRP 4

RESULT 14
 ID US-09-157-217-14366 STANDARD: PRI: 43 AA.
 XX XXXXXX

Sequence 14, Application US/0905051
 GENERAL INFORMATION:
 APPLICANT: Dr. Peter W. D.
 TITLE OF INVENTION: ISOLATED HUMAN PROTEIN NUCLEIC
 TITLE OF INVENTION: ANTISENSE MOLECULES ENCODING PROSOPIA PROTEINS, AND USES
 FILE REFERENCE: C1000162
 CURRENT FILING DATE: 1999-11-24
 NUMBER OF SEQ ID NOS: 33195
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 14366
 LENGTH: 43
 TYPE: PRT
 ORGANISM: DIOSOPHILA

Query Match: 100.0% Score 31 DB 2 Length 43.
 Best Local Similarity 100.0% Prod No. 2598-03
 Matches 4 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 24 VWRP 33
 QY 1 VWRP 4

RESULT 15
 ID US-09-095-051-240 STANDARD: PRI: 44 AA.
 XX XXXXXX

Sequence 240, Application US/0905051
 GENERAL INFORMATION:
 APPLICANT: Elaris Technology, Inc.
 APPLICANT: Duelett, Aymeric
 APPLICANT: DUELETT, AYMERIC
 TITLE OF INVENTION: 51 ESTS FOR SECRETED PROTEINS
 TITLE OF INVENTION: EXPRESSED IN VARIOUS TISSUES
 NUMBER OF SEQUENCES: 333
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knudsen, Morten, Olson & Bear
 STREET: 301 West Broadway
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92101-9005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY Disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: Win95
 SOFTWARE: Word
 CURRENT APPLICATION DATA: 2598-03
 APPLICATION NUMBER: 2598-03



CC APPLICANT: HEARING, Vincent J. 27.
 CC TITLE OF INVENTION: ISOLATED 2-PROTEIN COUPLED RECEPTORS.
 CC TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FOR PRIMING ANT-BODY
 CC TITLE OF INVENTION: THEREOF
 CC FILE REFERENCES: 0000174
 CC CURRENT FILING DATE: 1999-12-29
 CC CURRENT FILING DATE: 1999-12-29
 CC NUMBER OF SEQ. ID NOS: 2586
 CC SOFTWARE: FASTSEQ for Windows Version 4.0
 CC SEQ ID NO: 10074
 CC LENGTH: 59
 CC TYPE: PRT
 CC ORGANISM: Drosophila
 CC OTHER INFORMATION: Xaa means any amino acid
 CC SEQUENCE: 59 AA: 6068 MW: 20586 CN:

Query Match: 100.0% Score 39; DR 2; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6,146-02;
 Matches: 5; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 Db 42 VWRFP 46
 QY 1 VWRFP 5

RESULT 4
 ID US-09-174-044-0078 STANDARD: PRT: 59 AA.
 XX AC XXXXXX
 XX DI
 XX SEQUENCE 51160; Application US/09-174-0078
 CC GENERAL INFORMATION:
 CC APPLICANT: HEMINGER, et al.
 CC TITLE OF INVENTION: The Drosophila Genome: Primary Nucleic
 CC TITLE OF INVENTION: Acid Sequences of Genes and Transcripts Provided As
 CC TITLE OF INVENTION: Part of a Reference: Nucleic Acid Arrays and Kits
 CC FILE REFERENCES: 0000175
 CC CURRENT FILING DATE: 1999-12-28
 CC NUMBER OF SEQ. ID NOS: 42648
 CC SOFTWARE: FASTSEQ for Windows Version 4.0
 CC SEQ ID NO: 25078
 CC LENGTH: 59
 CC TYPE: PRT
 CC ORGANISM: Drosophila
 CC OTHER INFORMATION: Xaa means any amino acid
 CC SEQUENCE: 59 AA: 6068 MW: 20586 CN:

Query Match: 100.0% Score 39; DR 2; Length 59;
 Best Local Similarity 100.0%; Pred. No. 6,146-02;
 Matches: 5; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 Db 42 VWRFP 46
 QY 1 VWRFP 5

RESULT 5
 ID US-09-174-044-0078 STANDARD: PRT: 72 AA.
 XX AC XXXXXX
 XX DI
 XX SEQUENCE 51160; Application US/09-174-0078
 CC GENERAL INFORMATION:
 CC APPLICANT: HEARING, Vincent J. 27.
 CC TITLE OF INVENTION: ISOLATED 2-PROTEIN COUPLED RECEPTORS.
 CC TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FOR PRIMING ANT-BODY
 CC TITLE OF INVENTION: THEREOF
 CC FILE REFERENCES: 0000174
 CC CURRENT FILING DATE: 1999-12-29
 CC CURRENT FILING DATE: 1999-12-29
 CC NUMBER OF SEQ. ID NOS: 2586
 CC SOFTWARE: FASTSEQ for Windows Version 4.0
 CC SEQ ID NO: 10074
 CC LENGTH: 59
 CC TYPE: PRT
 CC ORGANISM: Drosophila
 CC OTHER INFORMATION: Xaa means any amino acid
 CC SEQUENCE: 59 AA: 6068 MW: 20586 CN:

CC CURRENT APPLICATION NUMBER: US/09-174-0078
 CC CURRENT FILING DATE: 1999-12-29
 CC NUMBER OF SEQ. ID NOS: 2586
 CC SOFTWARE: FASTSEQ for Windows Version 4.0
 CC SEQ ID NO: 3594
 CC LENGTH: 72
 CC TYPE: PRT
 CC ORGANISM: Drosophila melanogaster
 CC FEATURE:
 CC OTHER INFORMATION: Xaa means any amino acid
 CC SEQUENCE: 72 AA: 9528 MW: 24807 CN:
 Query Match: 100.0% Score 49; DR 17; Length 72;
 Best Local Similarity 100.0%; Pred. No. 6,146-02;
 Matches: 5; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 Db 42 VWRFP 46
 QY 1 VWRFP 5

RESULT 6
 ID US-09-270-767-0116 STANDARD: PRT: 72 AA.
 XX AC XXXXXX
 XX DI
 XX SEQUENCE 51160; Application US/09-270-767
 CC GENERAL INFORMATION:
 CC APPLICANT: HEMINGER, et al.
 CC TITLE OF INVENTION: Nucleic acids and proteins for a protein reference
 CC FILE REFERENCES: 0000174
 CC CURRENT FILING DATE: 1999-12-28
 CC NUMBER OF SEQ. ID NOS: 42648
 CC SOFTWARE: FASTSEQ for Windows Version 4.0
 CC SEQ ID NO: 5116
 CC LENGTH: 72
 CC TYPE: PRT
 CC ORGANISM: Drosophila melanogaster
 CC FEATURE:
 CC OTHER INFORMATION: Xaa means any amino acid
 CC SEQUENCE: 72 AA: 9528 MW: 24807 CN:

Query Match: 100.0% Score 49; DR 17; Length 72;
 Best Local Similarity 100.0%; Pred. No. 6,146-02;
 Matches: 5; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;
 Db 42 VWRFP 46
 QY 1 VWRFP 5

RESULT 7
 ID US-09-270-767-0116 STANDARD: PRT: 72 AA.
 XX AC XXXXXX
 XX DI
 XX SEQUENCE 51160; Application US/09-270-767
 CC GENERAL INFORMATION:
 CC APPLICANT: HEARING, Vincent J. 27.
 CC TITLE OF INVENTION: Nucleic acids and proteins for a protein reference
 CC TITLE OF INVENTION: THEREOF
 CC FILE REFERENCES: 0000174
 CC CURRENT FILING DATE: 1999-12-28
 CC NUMBER OF SEQ. ID NOS: 42648
 CC SOFTWARE: FASTSEQ for Windows Version 4.0
 CC SEQ ID NO: 5116
 CC LENGTH: 72
 CC TYPE: PRT
 CC ORGANISM: Drosophila melanogaster
 CC FEATURE:
 CC OTHER INFORMATION: Xaa means any amino acid
 CC SEQUENCE: 72 AA: 9528 MW: 24807 CN:

CC SIFTET, Suite 1200, 127 Peachtree St. N.E.
 CC CITY: Atlanta
 CC STATE: GA
 CC COUNTRY: USA
 CC ZIP: 30303-1811
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS, MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US 09/132-799
 CC FILING DATE: 22 Dec 1997
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 607,000,436
 CC FILING DATE: 22 Jun 1995
 CC CLASSIFICATION:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Spatt, Gerald D.
 CC REGISTRATION NUMBER: 36,016
 CC REFERENCE/DOCKET NUMBER: 14014, 0195
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (404) 688-9880
 CC TELEFAX: (404) 688-9880
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 131 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE: 131 AA; 14256 MW; 84197 CN;

Query Match: 100.0% Score 39, DB 14, Length 131;
 Best local similarity 100.0% Pred No. 6 146-02;
 Matches: 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 83 VWRPP 87
 QY 1 VWRPP 5
 |||||

RESULT 8
 ID US-09-083-132-2 STANDARD PRT: 131 AA
 XX xxxxxx
 AC xxxxxx
 DI |||||
 DT 1 VWRPP 5
 DE Sequence 2, Application US/0835517
 XX Sequence 2, Application US/0835517
 CC GENERAL INFORMATION:
 CC APPLICANT: HEARING, Vincent J., Jr.
 CC TITLE OF INVENTION: Depigmenting Activity of Aspartic Signal
 CC NUMBER OF SEQUENCES: 29
 CC CORRESPONDENCE ADDRESS:
 CC SIFTET, Suite 1200, 127 Peachtree St. N.E.
 CC CITY: Atlanta
 CC STATE: GA
 CC COUNTRY: USA
 CC ZIP: 30303-1811
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS, MS-DOS

CC SIFTET, Suite 1200, 127 Peachtree St. N.E.
 CC CITY: Atlanta
 CC STATE: GA
 CC COUNTRY: USA
 CC ZIP: 30303-1811
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS, MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US 09/132-799
 CC FILING DATE: 22 Dec 1997
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 607,000,436
 CC FILING DATE: 22 Jun 1995
 CC CLASSIFICATION:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Spatt, Gerald D.
 CC REGISTRATION NUMBER: 36,016
 CC REFERENCE/DOCKET NUMBER: 14014, 0195
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (404) 688-9880
 CC TELEFAX: (404) 688-9880
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 131 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE: 131 AA; 14256 MW; 84197 CN;

Query Match: 100.0% Score 39, DB 14, Length 131;
 Best local similarity 100.0% Pred No. 6 146-02;
 Matches: 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 83 VWRPP 87
 QY 1 VWRPP 5
 |||||

RESULT 9
 ID US-09-132-799-2 STANDARD PRT: 141 AA
 XX xxxxxx
 AC xxxxxx
 DI |||||
 DT 1 VWRPP 5
 DE Sequence 22215, Application US/09270267
 XX Sequence 22215, Application US/09270267
 CC GENERAL INFORMATION:
 CC APPLICANT: Homburger et al.
 CC TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 CC FILE REFERENCE: File Reference: 7326-094
 CC CURRENT APPLICATION NUMBER: US 09/132-799
 CC CURRENT FILING DATE: 1999-03-17
 CC NUMBER OF SEQ ID NOS: 21517
 CC SOFTWARE: Patent In Ver. 2.0
 CC SEQ ID NO: 13310
 CC LENGTH: 141
 CC TYPE: PRT
 CC ORGANISM: *Drosophila melanogaster*
 CC SEQUENCE: 141 AA; 15334 MW; 91905 CN;

Query Match: 100.0% Score 39, DB 17, Length 141;
 Best local similarity 100.0% Pred No. 6 146-02;
 Matches: 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

DB 81 VWRPP 85
 QY 1 VWRPP 5
 |||||

RESULT 10
 ID US-09-132-799-2 STANDARD PRT: 144 AA
 XX xxxxxx
 AC xxxxxx
 DI |||||
 DT 1 VWRPP 5
 DE Sequence 22215, Application US/09270267
 XX Sequence 22215, Application US/09270267
 CC GENERAL INFORMATION:
 CC APPLICANT: Homburger et al.
 CC TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 CC FILE REFERENCE: File Reference: 7326-094
 CC CURRENT APPLICATION NUMBER: US 09/132-799
 CC CURRENT FILING DATE: 1999-03-17
 CC NUMBER OF SEQ ID NOS: 21517
 CC SOFTWARE: Patent In Ver. 2.0
 CC SEQ ID NO: 13310
 CC LENGTH: 141
 CC TYPE: PRT
 CC ORGANISM: *Drosophila melanogaster*
 CC SEQUENCE: 144 AA; 15334 MW; 91905 CN;

Query Match: 100.0% Score 39, DB 17, Length 141;
 Best local similarity 100.0% Pred No. 6 146-02;
 Matches: 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

CC TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
 CC TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
 CC TITLE OF INVENTION: THEREOF

CC FILE REFERENCE: CL000173
 CC CURRENT APPLICATION NUMBER: US/60/173,404
 CC CURRENT FILING DATE: 1999-12-29
 CC NUMBER OF SEQ ID NOS: 30269
 CC SOFTWARE: FastSeq for Windows Version 4.0
 CC SEQ ID NO 12740
 CC LENGTH: 196
 CC TYPE: PRI
 CC ORGANISM: Drosophila
 CC SEQUENCE 196 AA: 22128 MW: 190309 CN:

Query Match: 100.0%, Score 39, DB 2, Length 196;
 Best Local Similarity 100.0%; Pred. No: 6,14e-02;
 Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

Db 188 VWRPP 192
 |||||
 QV 1 VWRPP 5

RESULT 15

ID US-60-157-217-13541 STANDARD; PRI: 196 AA.

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

XX XXXXX

Sequence 15541, Application US/50167217

Sequence 15541, Application US/50167217

GENERAL INFORMATION:

APPLICANT: LI, Peter W. D.

TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING DROSOPHILA PROTEINS, AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL000152

CURRENT APPLICATION NUMBER: US/60/173,404

CURRENT FILING DATE: 1999-11-24

NUMBER OF SEQ ID NOS: 23195

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15541

LENGTH: 196

TYPE: PRI

ORGANISM: Drosophila

SEQUENCE 195 AA: 22128 MW: 190309 CN:

Query Match: 100.0%, Score 39, DB 2, Length 196;
 Best Local Similarity 100.0%; Pred. No: 6,14e-02;
 Matches 5, Conservative 0, Mismatches 0, Indels 0, Gaps 0,

Db 188 VWRPP 192
 |||||
 QV 1 VWRPP 5

Search Completed: Fri Feb 25 13:23:49 2000
 Job time : 229 secs.

CC FILING DATE: 23 JUN 1995
 CC CLASSIFICATION: S14
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Spratt, Gwendolyn D.
 CC REGISTRATION NUMBER: 36,016
 CC REFERENCE/KEY NUMBER: 14014 01-5
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (404) 688-0770
 CC TELEFAX: (404) 688-9880
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 131 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE: 131 AA: 14256 MW: 84197 CN:

Query Match: 100.0% Score 47 DB 14 Length 131
 Best Local Similarity 100.0% Prod No 1 96e-02
 Matches: 6 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 83 VRPPP 88
 QY 1 VRPPP 6

RESULT 2
 ID US-60-000-436-2 STANDARD PPT: 131 AA

CC xxxxxx

Sequence 2 Application US/60000436

SCHEME 2 Application US/60000436
 GENERAL INFORMATION:
 CC APPLICANT: HEALING Vincent J. Jr.
 CC TITLE OF INVENTION: Depigmentation Activity of Aquatic Signal
 CC TITLE OF INVENTION: Protein and Peptides Thereof
 CC NUMBER OF SEQUENCES: 29
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Needle & Posenberry, P.C.
 CC STREET: Suite 1200, 127 Peachtree St., N.E.
 CC CITY: Atlanta
 CC STATE: GA
 CC COUNTRY: USA
 CC ZIP: 30303-1811

COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/60/000,436
 CC FILING DATE: 22 Dec 1997

CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US96/10695
 CC FILING DATE: 21 Jun 1996

CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 602,000,436
 CC FILING DATE: 23 Jun 1995

CLASSIFICATION:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Spratt, Gwendolyn D.
 CC REGISTRATION NUMBER: 36,016
 CC REFERENCE/KEY NUMBER: 14014 01-5
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (404) 688-0770
 CC TELEFAX: (404) 688-9880

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 131 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE: 131 AA: 14256 MW: 84197 CN:

Query Match: 100.0% Score 47 DB 20 Length 131
 Best Local Similarity 100.0% Prod No 1 96e-02
 Matches: 6 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 83 VRPPP 88
 QY 1 VRPPP 6

RESULT 4
 ID US-60-417-507-20071 STANDARD PPT: 70 AA

CC xxxxxx

Sequence 20071 Application US/00417507

SCHEME 22971 Application US/00417507
 GENERAL INFORMATION:
 CC APPLICANT: KEITH G. WEINSTOCK ET AL.
 CC TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILL
 CC TITLE OF INVENTION: FUNGICIDES FOR DIAGNOSTICS AND THERAPEUTICS
 CC FILE REFERENCE: PATH99-10
 CC CURRENT APPLICATION NUMBER: US/00/417,507
 CC CURRENT FILING DATE: 1999-10-14
 CC SEQ ID NO 20071
 CC LENGTH: 70
 CC TYPE: PPT
 CC ORGANISM: A.fumigatus

Query Match: 95.7% Score 45 DB 19 Length 70
 Best Local Similarity 93.0% Prod No 3 32e-02
 Matches: 5 Conservative 1 Mismatches 0 Indels 0 Gaps 0

DB 19 VRPPP 24
 QY 1 VRPPP 6

Sequence 30179 Application US/50173464

Query Match: 95.7% Score 45 DB 19 Length 70
 Best Local Similarity 93.0% Prod No 3 32e-02
 Matches: 5 Conservative 1 Mismatches 0 Indels 0 Gaps 0

RESULT 5

ID US-60-174-464-40179 STANDARD PPT: 216 AA

CC xxxxxx

Sequence 30179 Application US/50173464

SCHEME 30179 Application US/50173464
 GENERAL INFORMATION:
 CC APPLICANT: LI, Peter W.D.
 CC TITLE OF INVENTION: ISOLATED PROTEIN LIGAND RECEPTORS,
 CC TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING SUCH PROTEINS AND USES
 CC TITLE OF INVENTION: THESEOF
 CC FILE REFERENCE: 01090173

Query Match: 95.7% Score 45 DB 19 Length 70
 Best Local Similarity 93.0% Prod No 3 32e-02
 Matches: 5 Conservative 1 Mismatches 0 Indels 0 Gaps 0

DB 19 VRPPP 24
 QY 1 VRPPP 6

[illegible][illegible][illegible]

CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1 0, Version #1 25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/99/000,572A
 CC FILING DATE: 12-APR-1996
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Campbell, Cathryn A.
 CC REGISTRATION NUMBER: 31,815
 CC REFERENCE/DOCKET NUMBER: P-PM 1997
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (619) 535-9001
 CC TELEFAX: (619) 535-8949
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 532 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE: 532 AA: 54263 MW: 1541840 CN:

Query Match 95.7% Score 45. DB 11. Length 532:
 Best Local Similarity 83.3% Pred. No. 3,329-02:
 Matches 5. Conservative 1. Mismatches 0. Indels 0. Gaps 0.

Db 36 VIRPPP 41
 1:111
 QY 1 WRPPP 6

RESULT 10
 ID US-09-450-669-3924 STANDARD: PRI: 198 AA.

XX
 AC xxxxxx

DI
 XX

DE Sequence 3924, Application US/99/450669

XX Sequence 3924, Application US/99/450669

CC GENERAL INFORMATION:

CC APPLICANT: Lynn Doucette Stamm et al.

CC TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO SIAPHYLMOO

CC TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

CC FILE REFERENCE: PATH99-09A

CC CURRENT APPLICATION NUMBER: US/99/450 669

CC CURRENT FILING DATE: 1999-11-29

CC NUMBER OF SEQ ID NOS: 7544

CC SEQ ID NO 3924

LENGTH: 198

TYPE: PRI

CC ORGANISM: S.epidermidis

CC SEQUENCE: 198 AA: 21600 MW: 222826 CN:

Query Match 91.5% Score 43. DB 19. Length 198:

Best Local Similarity 66.7% Pred No. 5,879-02:

Matches 12. Conservative 2. Mismatches 0. Indels 0. Gaps 0.

Db 127 IIRPPP 132

1:111

QY 1 WRPPP 6

Search completed: Fri Feb 25 13:34:12 2000

Job time : 2.7 secs.


```

383 20 03 0 1107 1 849660 1240+03
384 20 03 0 1107 1 849660 1240+03
385 20 03 0 1107 1 849660 1240+03
386 20 03 0 1107 1 849660 1240+03
387 20 03 0 1107 1 849660 1240+03
388 20 03 0 1107 1 849660 1240+03
389 20 03 0 1107 1 849660 1240+03
390 20 03 0 1107 1 849660 1240+03
391 20 03 0 1107 1 849660 1240+03
392 20 03 0 1107 1 849660 1240+03
393 20 03 0 1107 1 849660 1240+03

```

Nucleotide sequence of random 607 sequences from list due to search parameters chosen

ALIGNMENTS

```

RESULT 1
ID R56726 Standard: peptide: 12 AA.
AC R56726: (first entry)
DE 20-MAR-1996
FW Epitope: B cell epitope isolated by anti-lymphoma E AL panel.
KW Epitope: B cell epitope isolated by anti-lymphoma E AL panel.
OS Synthetic
FH Key
FT region
FT 419
FT 1107-12 1107-12
PN US9338665-A.
PD 15-AUG-1994.
FE 15-OCT-1991: US-778233.
PR 15-OCT-1991: US-778233.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Schatz PJ, Stemmer WPC.
PI WPI 94-263274/32.
PT Construction of random peptide library - by creating vectors
PT containing DNA encoding the random peptide(s) fused to DNA binding
PT proteins; used to screen for novel ligands
PS Example 4: Fig 3A: 49pp: English.
CC A random peptide library was constructed in E coli hosts
CC the library was used to screen for novel ligands which
CC recognize the tyrosinase epitope RARVW. Peptides isolated by
CC panning were sequenced and a consensus epitope was identified (see
CC features table). Arginine is invariant in the first position for all
CC the virus positive clones (residues 1-5). A strong bias was
CC evident for the second position but in the third position, 5 amino
CC acids (His, Asp, Tyr, Trp) account for 70% of the residues. The
CC fourth position shows a strong bias for positively charged residues
CC (Lys and Arg) with almost exclusively hydrophobic residues at
CC position 7 (mostly Val). Val and Thr predominated at the sixth
CC position (70%) with Ser and His accounting for the remaining amino
CC acids.
SQ Sequence 12 AA.

```

```

Query Match: 100.0%, Score 31, PP 12, Length 12.
Best local similarity: 100.0%, Pred No 5, 57+02.
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0.
DB 8 vrp 11
QY 1 VWP 4

```

```

RESULT 2
ID R56726 Standard: peptide: 12 AA.
AC R56726: (first entry)
DE 20-MAR-1996
FW Epitope: B cell epitope isolated by anti-lymphoma E AL panel.
KW Epitope: B cell epitope isolated by anti-lymphoma E AL panel.
OS Synthetic
FH Key
FT region
FT 419
FT 1107-12 1107-12
PN US9338665-A.
PD 15-AUG-1994.
FE 15-OCT-1991: US-778233.
PR 15-OCT-1991: US-778233.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Schatz PJ, Stemmer WPC.
PI WPI 94-263274/32.
PT Construction of random peptide library - by creating vectors
PT containing DNA encoding the random peptide(s) fused to DNA binding
PT proteins; used to screen for novel ligands
PS Example 4: Fig 3A: 49pp: English.
CC A random peptide library was constructed in E coli hosts
CC the library was used to screen for novel ligands which
CC recognize the tyrosinase epitope RARVW. Peptides isolated by
CC panning were sequenced and a consensus epitope was identified (see
CC features table). Arginine is invariant in the first position for all
CC the virus positive clones (residues 1-5). A strong bias was
CC evident for the second position but in the third position, 5 amino
CC acids (His, Asp, Tyr, Trp) account for 70% of the residues. The
CC fourth position shows a strong bias for positively charged residues
CC (Lys and Arg) with almost exclusively hydrophobic residues at
CC position 7 (mostly Val). Val and Thr predominated at the sixth
CC position (70%) with Ser and His accounting for the remaining amino
CC acids.
SQ Sequence 12 AA.

```

```

FH Key
FT region
FT 419
FT 1107-12 1107-12
PN US9338665-A.
PD 15-AUG-1994.
FE 15-OCT-1991: US-778233.
PR 15-OCT-1991: US-778233.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Schatz PJ, Stemmer WPC.
PI WPI 94-263274/32.
PT Construction of random peptide library - by creating vectors
PT containing DNA encoding the random peptide(s) fused to DNA binding
PT proteins; used to screen for novel ligands
PS Example 4: Fig 3A: 49pp: English.
CC A random peptide (PP) library can be constructed by transforming host
CC cells with a collection of a DNA recombinant vectors that encode a fusion
CC protein comprised of a DNA binding protein (BP) and a PP and also
CC contains a binding site for the BP. The PP library can be used to
CC screen for novel ligands, the method resulting in the formation of a
CC complex comprising the fusion protein bound to a receptor through the BP
CC ligand and to the recombinant DNA vector through the BP. An EP
CC library (Epitope) was screened with RARVW and a six amino acid region
CC of tyrosinase (RARVW) as a target peptide. An EP library was found to the preferred
CC recognition sequence for RARVW.
SQ Sequence 12 AA.

```

```

Query Match: 100.0%, Score 31, PP 12, Length 12.
Best local similarity: 100.0%, Pred No 5, 57+02.
Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0.
DB 8 vrp 11
QY 1 VWP 4

```

```

RESULT 3
ID W25566 Standard: peptide: 12 AA.
AC W25566: (first entry)
DE 14-OCT-1997
DE Activity: 12 12 12 12 12 12 12 12 12 12 12 12
FW Random peptide library, DNA binding protein, receptor ligand, linker:
KW Random peptide library, DNA binding protein, receptor ligand, linker:
OS Synthetic.
FH Key
FT region
FT 419
FT 1107-12 1107-12
PN W25566: (first entry)
PD 14-OCT-1997.
FE 14-OCT-1997.
PR 14-OCT-1997.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
PI Cull MG, Gates CM, Miller JE, Schatz PJ, Stemmer WPC.
PI WPI 94-263274/32.
PT Random peptide library and affinity enrichment methods for screening
PT it is useful to identify peptides that bind receptor proteins, of
PT interest, useful for therapeutic, diagnostic and related purposes
PS Example 4: Fig 3A: 49pp: English.
CC W25566 represent epitopes for the antibody D22.39. These sequences
CC were isolated by a method of the invention to isolate a DNA binding
CC protein, or a peptide with specific affinity for a receptor. The method
CC comprises providing a recombinant DNA vector encoding a peptide having
CC specific affinity for a receptor. A library of oligonucleotides encoding
CC different potential DNA binding proteins is inserted into the vector. The
CC vector is transformed into a fusion protein. A fusion protein comprises
CC a potential DNA binding protein with affinity for the receptor. The fusion
CC protein binds to the receptor to form a complex. The host cells are lysed
CC to isolate the complex which are screened with a receptor to induce
CC peptide binding to the receptor. The random peptide library and the

```

CC methods for screening it can be used to identify peptides that bind
 CC receptor molecules of interest. The peptides can be used for therapeutic,
 CC diagnostic and related purposes, e.g., to bind the receptor, or an
 CC antagonist, and so inhibit or promote the activity of the receptor. The
 CC method of affinity enrichment allows a very large library of peptides to
 CC be screened, and by identifying the peptide de novo, the sequence or
 CC structure of the receptor molecule or the natural binding partner of the
 CC receptor need not be known.

CC Sequence 12 AA:

Query Match: 100.0% Score 31; IP 27; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5,676-02;

Matches 4 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 8 vwp 11

QY 1 VWP 4

RESULT 5

ID W12207 standard; Peptide: 12 AA;
 AC W12207;

DI 25-MAR-1997 (first entry)

DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response

OS Synthetic

PN US5578496-A

PD 26-NOV-1994

PF 19-DEC-1991; 81106

PR 19-DEC-1991; US-811159

PR 20-OCT-1992; US-965670

PR 05-OCT-1994; US-518200

PA (RAYU) HAYLER COLLEGE MEDICINE

PI Ashlunda, T. Atassi, MZ

WP: 97-02122/22

DI 25-MAR-1997 (first entry)

DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response

OS Synthetic

PN US5578496-A

PD 26-NOV-1994

PF 19-DEC-1991; 81106

PR 19-DEC-1991; US-811159

PR 20-OCT-1992; US-965670

PR 05-OCT-1994; US-518200

PA (RAYU) HAYLER COLLEGE MEDICINE

PI Ashlunda, T. Atassi, MZ

WP: 97-02122/22

DI 25-MAR-1997 (first entry)

DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response

OS Synthetic

PN US5578496-A

PD 26-NOV-1994

PF 19-DEC-1991; 81106

PR 19-DEC-1991; US-811159

PR 20-OCT-1992; US-965670

PR 05-OCT-1994; US-518200

PA (RAYU) HAYLER COLLEGE MEDICINE

PI Ashlunda, T. Atassi, MZ

WP: 97-02122/22

DI 25-MAR-1997 (first entry)

DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response

DE Human acetylcholine receptor alpha chain amino acids 14-27.
 KW Tetracycline compound, derivatization polymer: polystyrene-lysine
 KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease
 KW Graves disease, antibody, antibody response
 OS Synthetic
 PN US5578496-A
 PD 26-NOV-1994
 PF 19-DEC-1991; 81106
 PR 19-DEC-1991; US-811159
 PR 20-OCT-1992; US-965670
 PR 05-OCT-1994; US-518200
 PA (RAYU) HAYLER COLLEGE MEDICINE
 PI Ashlunda, T. Atassi, MZ
 WP: 97-02122/22

Query Match: 100.0% Score 31; IP 27; Length 12;

Best Local Similarity 100.0%; Pred. No. 6,426-02;

Matches 4 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 8 vwp 11

QY 1 VWP 4

RESULT 5

ID W12207 standard; Peptide: 16 AA;
 AC W12207;

DI 25-MAR-1997 (first entry)

DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response

OS Synthetic

PN US5578496-A

PD 26-NOV-1994

PF 19-DEC-1991; 81106

PR 19-DEC-1991; US-811159

PR 20-OCT-1992; US-965670

PR 05-OCT-1994; US-518200

PA (RAYU) HAYLER COLLEGE MEDICINE

PI Ashlunda, T. Atassi, MZ

WP: 97-02122/22

DI 25-MAR-1997 (first entry)

DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response

OS Synthetic

PN US5578496-A

PD 26-NOV-1994

PF 19-DEC-1991; 81106

PR 19-DEC-1991; US-811159

PR 20-OCT-1992; US-965670

PR 05-OCT-1994; US-518200

PA (RAYU) HAYLER COLLEGE MEDICINE

PI Ashlunda, T. Atassi, MZ

WP: 97-02122/22

DI 25-MAR-1997 (first entry)







DE Human acetylcholine receptor alpha chain amino acids 14-27.

KW Tetracycline compound, derivatization polymer: polystyrene-lysine

KW polyvinyl alcohol, immunotherapy, myasthenia gravis, autoimmune disease

KW Graves disease, antibody, antibody response



Revised by A. J. A. Collins, Biocomputing Research Unit,
Department of Chemistry, University of Edinburgh, J.K.
Distribution rights by Oxford Molecular Ltd

Waterfall plot for protein-protein database search, using Smith-Waterman algorithm
 Run on: Fri Feb 25 14:06:53 2006; Master time 1.67 Seconds
 Labeled output not generated. 31.869 Million cell updates/sec

1. $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$
 2. $\mathcal{B} = \{B_1, B_2, \dots, B_m\}$
 3. $\mathcal{C} = \{C_1, C_2, \dots, C_k\}$
 4. $\mathcal{D} = \{D_1, D_2, \dots, D_l\}$
 5. $\mathcal{E} = \{E_1, E_2, \dots, E_p\}$
 6. $\mathcal{F} = \{F_1, F_2, \dots, F_q\}$
 7. $\mathcal{G} = \{G_1, G_2, \dots, G_r\}$
 8. $\mathcal{H} = \{H_1, H_2, \dots, H_s\}$
 9. $\mathcal{I} = \{I_1, I_2, \dots, I_t\}$
 10. $\mathcal{J} = \{J_1, J_2, \dots, J_u\}$
 11. $\mathcal{K} = \{K_1, K_2, \dots, K_v\}$
 12. $\mathcal{L} = \{L_1, L_2, \dots, L_w\}$
 13. $\mathcal{M} = \{M_1, M_2, \dots, M_x\}$
 14. $\mathcal{N} = \{N_1, N_2, \dots, N_y\}$
 15. $\mathcal{O} = \{O_1, O_2, \dots, O_z\}$
 16. $\mathcal{P} = \{P_1, P_2, \dots, P_{10}\}$
 17. $\mathcal{Q} = \{Q_1, Q_2, \dots, Q_{10}\}$
 18. $\mathcal{R} = \{R_1, R_2, \dots, R_{10}\}$
 19. $\mathcal{S} = \{S_1, S_2, \dots, S_{10}\}$
 20. $\mathcal{T} = \{T_1, T_2, \dots, T_{10}\}$
 21. $\mathcal{U} = \{U_1, U_2, \dots, U_{10}\}$
 22. $\mathcal{V} = \{V_1, V_2, \dots, V_{10}\}$
 23. $\mathcal{W} = \{W_1, W_2, \dots, W_{10}\}$
 24. $\mathcal{X} = \{X_1, X_2, \dots, X_{10}\}$
 25. $\mathcal{Y} = \{Y_1, Y_2, \dots, Y_{10}\}$
 26. $\mathcal{Z} = \{Z_1, Z_2, \dots, Z_{10}\}$
 27. $\mathcal{AA} = \{A_1, A_2, \dots, A_{10}\}$
 28. $\mathcal{BB} = \{B_1, B_2, \dots, B_{10}\}$
 29. $\mathcal{CC} = \{C_1, C_2, \dots, C_{10}\}$
 30. $\mathcal{DD} = \{D_1, D_2, \dots, D_{10}\}$
 31. $\mathcal{EE} = \{E_1, E_2, \dots, E_{10}\}$
 32. $\mathcal{FF} = \{F_1, F_2, \dots, F_{10}\}$
 33. $\mathcal{GG} = \{G_1, G_2, \dots, G_{10}\}$
 34. $\mathcal{HH} = \{H_1, H_2, \dots, H_{10}\}$
 35. $\mathcal{II} = \{I_1, I_2, \dots, I_{10}\}$
 36. $\mathcal{JJ} = \{J_1, J_2, \dots, J_{10}\}$
 37. $\mathcal{KK} = \{K_1, K_2, \dots, K_{10}\}$
 38. $\mathcal{LL} = \{L_1, L_2, \dots, L_{10}\}$
 39. $\mathcal{MM} = \{M_1, M_2, \dots, M_{10}\}$
 40. $\mathcal{NN} = \{N_1, N_2, \dots, N_{10}\}$
 41. $\mathcal{OO} = \{O_1, O_2, \dots, O_{10}\}$
 42. $\mathcal{PP} = \{P_1, P_2, \dots, P_{10}\}$
 43. $\mathcal{QQ} = \{Q_1, Q_2, \dots, Q_{10}\}$
 44. $\mathcal{RR} = \{R_1, R_2, \dots, R_{10}\}$
 45. $\mathcal{SS} = \{S_1, S_2, \dots, S_{10}\}$
 46. $\mathcal{TT} = \{T_1, T_2, \dots, T_{10}\}$
 47. $\mathcal{UU} = \{U_1, U_2, \dots, U_{10}\}$
 48. $\mathcal{VV} = \{V_1, V_2, \dots, V_{10}\}$
 49. $\mathcal{WW} = \{W_1, W_2, \dots, W_{10}\}$
 50. $\mathcal{XX} = \{X_1, X_2, \dots, X_{10}\}$
 51. $\mathcal{YY} = \{Y_1, Y_2, \dots, Y_{10}\}$
 52. $\mathcal{ZZ} = \{Z_1, Z_2, \dots, Z_{10}\}$
 53. $\mathcal{AAA} = \{A_1, A_2, \dots, A_{10}\}$
 54. $\mathcal{BBB} = \{B_1, B_2, \dots, B_{10}\}$
 55. $\mathcal{CCC} = \{C_1, C_2, \dots, C_{10}\}$
 56. $\mathcal{DDD} = \{D_1, D_2, \dots, D_{10}\}$
 57. $\mathcal{EEE} = \{E_1, E_2, \dots, E_{10}\}$
 58. $\mathcal{FFF} = \{F_1, F_2, \dots, F_{10}\}$
 59. $\mathcal{GGG} = \{G_1, G_2, \dots, G_{10}\}$
 60. $\mathcal{HHH} = \{H_1, H_2, \dots, H_{10}\}$
 61. $\mathcal{III} = \{I_1, I_2, \dots, I_{10}\}$
 62. $\mathcal{JJJ} = \{J_1, J_2, \dots, J_{10}\}$
 63. $\mathcal{KKK} = \{K_1, K_2, \dots, K_{10}\}$
 64. $\mathcal{LLL} = \{L_1, L_2, \dots, L_{10}\}$
 65. $\mathcal{MMM} = \{M_1, M_2, \dots, M_{10}\}$
 66. $\mathcal{NNN} = \{N_1, N_2, \dots, N_{10}\}$
 67. $\mathcal{OOO} = \{O_1, O_2, \dots, O_{10}\}$
 68. $\mathcal{PPP} = \{P_1, P_2, \dots, P_{10}\}$
 69. $\mathcal{QQQ} = \{Q_1, Q_2, \dots, Q_{10}\}$
 70. $\mathcal{RRR} = \{R_1, R_2, \dots, R_{10}\}$
 71. $\mathcal{SSS} = \{S_1, S_2, \dots, S_{10}\}$
 72. $\mathcal{TTT} = \{T_1, T_2, \dots, T_{10}\}$
 73. $\mathcal{UUU} = \{U_1, U_2, \dots, U_{10}\}$
 74. $\mathcal{VVV} = \{V_1, V_2, \dots, V_{10}\}$
 75. $\mathcal{WWW} = \{W_1, W_2, \dots, W_{10}\}$
 76. $\mathcal{XXX} = \{X_1, X_2, \dots, X_{10}\}$
 77. $\mathcal{YYY} = \{Y_1, Y_2, \dots, Y_{10}\}$
 78. $\mathcal{ZZZ} = \{Z_1, Z_2, \dots, Z_{10}\}$
 79. $\mathcal{AAAA} = \{A_1, A_2, \dots, A_{10}\}$
 80. $\mathcal{BBBB} = \{B_1, B_2, \dots, B_{10}\}$
 81. $\mathcal{CCCC} = \{C_1, C_2, \dots, C_{10}\}$
 82. $\mathcal{DDDD} = \{D_1, D_2, \dots, D_{10}\}$
 83. $\mathcal{EEEE} = \{E_1, E_2, \dots, E_{10}\}$
 84. $\mathcal{FFFF} = \{F_1, F_2, \dots, F_{10}\}$
 85. $\mathcal{GGGG} = \{G_1, G_2, \dots, G_{10}\}$
 86. $\mathcal{HHHH} = \{H_1, H_2, \dots, H_{10}\}$
 87. $\mathcal{IIII} = \{I_1, I_2, \dots, I_{10}\}$
 88. $\mathcal{JJJJ} = \{J_1, J_2, \dots, J_{10}\}$
 89. $\mathcal{KKKK} = \{K_1, K_2, \dots, K_{10}\}$
 90. $\mathcal{LLLL} = \{L_1, L_2, \dots, L_{10}\}$
 91. $\mathcal{MMMM} = \{M_1, M_2, \dots, M_{10}\}$
 92. $\mathcal{NNNN} = \{N_1, N_2, \dots, N_{10}\}$
 93. $\mathcal{OOOO} = \{O_1, O_2, \dots, O_{10}\}$
 94. $\mathcal{PPPP} = \{P_1, P_2, \dots, P_{10}\}$
 95. $\mathcal{QQQQ} = \{Q_1, Q_2, \dots, Q_{10}\}$
 96. $\mathcal{RRRR} = \{R_1, R_2, \dots, R_{10}\}$
 97. $\mathcal{SSSS} = \{S_1, S_2, \dots, S_{10}\}$
 98. $\mathcal{TTTT} = \{T_1, T_2, \dots, T_{10}\}$
 99. $\mathcal{UUUU} = \{U_1, U_2, \dots, U_{10}\}$
 100. $\mathcal{VVVV} = \{V_1, V_2, \dots, V_{10}\}$
 101. $\mathcal{WWWW} = \{W_1, W_2, \dots, W_{10}\}$
 102. $\mathcal{XXXX} = \{X_1, X_2, \dots, X_{10}\}$
 103. $\mathcal{YYYY} = \{Y_1, Y_2, \dots, Y_{10}\}$
 104. $\mathcal{ZZZZ} = \{Z_1, Z_2, \dots, Z_{10}\}$
 105. $\mathcal{AAAAA} = \{A_1, A_2, \dots, A_{10}\}$
 106. $\mathcal{BBBBB} = \{B_1, B_2, \dots, B_{10}\}$
 107. $\mathcal{CCCCC} = \{C_1, C_2, \dots, C_{10}\}$
 108. $\mathcal{DDDDD} = \{D_1, D_2,$

[illegible][illegible]

Score (1963) Mean 11.063; Variance 28.015; scale 0.427

[illegible][illegible]

[illegible]

CC APPLICATION NUMBER: US09/548,540
 CC FILING DATE: 15-SEP-1994
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/590,641
 CC FILING DATE: 15-AUG-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07,563,321
 CC FILING DATE: 15-OCT-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Smith, William M.
 CC REGISTRATION NUMBER: 30,223
 CC REFERENCE/DOCKET NUMBER: 11509-50-1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415-326-2400
 CC TELEFAX: 415-326-2422
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 12 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC IMMEDIATE SOURCE:
 CC CLONE: 69-2-1-1
 CC SEQUENCE 12 AA: 1376 MW: 912 CN:
 CC
 CC Query Match: 100.0% Score 31: PB 1: Length 12:
 CC Best Local Similarity 100.0% Pred No. 4,640-02:
 CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 CC
 Db 8 VWRP 11
 QY 1 VWRP 4
 CC
 CC RESULT 3
 CC ID US-09-132-799-232P-32 STANDARD: REF: 12 AA:
 CC AC xxxxxx
 CC
 CC Application 02 Application US/07799-232P
 CC DE US-09-132-799-232P-32 STANDARD: REF: 12 AA:
 CC Patent No. 5270170
 CC GENERAL INFORMATION:
 CC APPLICANT: Schatz, Peter J.
 CC APPLICANT: Cull, Millard G.
 CC APPLICANT: Miller, Jeff F.
 CC TITLE OF INVENTION: Peptide Library and Screening Method
 CC NUMBER OF SEQUENCES: 99
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESS: William M. Smith
 CC Street One Market Place, Ste 201, San Francisco, CA 94105
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 94105
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US 07,779,223
 CC FILING DATE: 19910106
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07,779,223
 CC FILING DATE: 15-OCT-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Smith, William M.
 CC REGISTRATION NUMBER: 30,223
 CC REFERENCE/DOCKET NUMBER: 11509-50-1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415-326-2400
 CC TELEFAX: 415-326-2422
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 12 amino acids
 CC TYPE: AMINO ACID
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear

CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415-326-2400
 CC TELEFAX: 415-326-2422
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 12 amino acids
 CC TYPE: AMINO ACID
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC IMMEDIATE SOURCE:
 CC CLONE: 69-2-1-1
 CC SEQUENCE 12 AA: 1376 MW: 912 CN:
 CC
 CC Query Match: 100.0% Score 31: PB 1: Length 12:
 CC Best Local Similarity 100.0% Pred No. 4,640-02:
 CC Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 CC
 Db 8 VWRP 11
 QY 1 VWRP 4
 CC
 CC RESULT 4
 CC ID US-09-132-799-232P-32 STANDARD: REF: 12 AA:
 CC AC xxxxxx
 CC
 CC Sequence 12, Application US/07799-232P
 CC Patent No. 5338665
 CC GENERAL INFORMATION:
 CC APPLICANT: Schatz, Peter J.
 CC APPLICANT: Cull, Millard G.
 CC APPLICANT: Miller, Jeff F.
 CC APPLICANT: Stemmer, Willem P.C.
 CC TITLE OF INVENTION: Peptide Library and Screening Method
 CC NUMBER OF SEQUENCES: 99
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: William M. Smith
 CC Street One Market Place, Ste 201, San Francisco, CA 94105
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: USA
 CC ZIP: 94105
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US 07,779,223
 CC FILING DATE: 19910115
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07,779,223
 CC FILING DATE: 15-OCT-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Smith, William M.
 CC REGISTRATION NUMBER: 30,223
 CC REFERENCE/DOCKET NUMBER: 11509-50-1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 415-326-2400
 CC TELEFAX: 415-326-2422
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 12 amino acids
 CC TYPE: AMINO ACID
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear

MATCHES	4:	Conservative		Mismatches		Gaps	
DB	R VWRP	11					
OY	I VWRP	4					
RESULT	6						
ID	US-08-31B-200-R	STANDARD	PRT:	10 AA			
XX	XXXXX						
AC							
XX							
DT							
XX							
DE	Sequence 6, Application US/080420						
CC	Sequence 6, Application US/080420						
CC	Patent No. 557447						
CC	GENERAL INFORMATION:						
CC	APPLICANT: ALABAMA, M. Z.						
CC	APPLICANT: ASHBAUGH, T.						
CC	TITLE INVENTION: N-TERMINAL PVA-EPIPEPTIDE WITH 199-N SULFHYDRYL						
CC	TITLE INVENTOR: DEPTHS FOR EPIPEPTIDES WITH 199-N SULFHYDRYL						
CC	NUMBER OF SEQUENCES: 25						
CC	CORRESPONDENCE ADDRESS:						
CC	ADDRESSEE: ARTHUR WHITE & BARKER						
CC	STREET: PO BOX 4449						
CC	CITY: HOUSTON						
CC	STATE: TEXAS						
CC	COUNTRY: USA						
CC	ZIP: 77210						
CC	COMPUTER READABLE FORM:						
CC	MEDIUM TYPE: floppy disk						
CC	COMPUTER: IBM PC compatible						
CC	OPERATING SYSTEM: DOS/MS-DOS						
CC	SOFTWARE: Patent Release #1.0, Version #1.25						
CC	CURRENT APPLICATION DATA:						
CC	APPLICATION NUMBER: US/8706420						
CC	FILING DATE:						
CC	CLASSIFICATION: 435						
CC	PRIOR APPLICATION DATA:						
CC	APPLICATION NUMBER: US/7965770						
CC	FILING DATE:						
CC	ATTORNEY/AGENT INFORMATION:						
CC	NAME: McDaniel, C. Steven						
CC	TELEPHONE: 713-787-1400						
CC	TELEPHONE: 713-787-1400						
CC	TELEPHONE: 713-787-1400						
CC	TELEPHONE: 713-787-1400						
CC	INFORMATION FOR SEQUENCE:						
CC	SEQUENCE CHARACTERISTICS:						
CC	LENGTH: 16 amino acids						
CC	TYPE: amino acid						
CC	TOP LOGS: linear						
CC	MOLECULE TYPE: polypeptide						
CC	SEQUENCE: 16 AA: 1062 MW: 1478 CN:						
CC	Quality Match:						
CC	Best local similarity 100%, Prod. No. 4640-21						
Matches	4:	Conservative		Mismatches		Gaps	
DB	7 VWRP	10					
OY	I VWRP	4					
RESULT	7						
ID	US-09-047-125-14	STANDARD	PRT:	92 AA			
XX	XXXXX						
AC							
XX							
DT							
XX							
DE	Sequence 14, Application US/8706420						

MATCHES	4:	Conservative		Mismatches		Gaps	
DB	R VWRP	11					
OY	I VWRP	4					
RESULT	6						
ID	US-08-31B-200-R	STANDARD	PRT:	10 AA			
XX	XXXXX						
AC							
XX							
DT							
XX							
DE	Sequence 6, Application US/080420						
CC	Sequence 6, Application US/080420						
CC	Patent No. 557447						
CC	GENERAL INFORMATION:						
CC	APPLICANT: ALABAMA, M. Z.						
CC	APPLICANT: ASHBAUGH, T.						
CC	TITLE INVENTION: N-TERMINAL PVA-EPIPEPTIDE WITH 199-N SULFHYDRYL						
CC	TITLE INVENTOR: DEPTHS FOR EPIPEPTIDES WITH 199-N SULFHYDRYL						
CC	NUMBER OF SEQUENCES: 25						
CC	CORRESPONDENCE ADDRESS:						
CC	ADDRESSEE: ARTHUR WHITE & BARKER						
CC	STREET: PO BOX 4449						
CC	CITY: HOUSTON			</			

XX Sequence 14, Application US/0852816A
 CC Patent No. 5872215
 CC GENERAL INFORMATION:
 CC APPLICANT: Leand F. Vellard, Peter Brimovskis,
 CC TITLE OF INVENTION: MARSH'S DISCOURSE REPRESENTATIONS
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Ian C. McLeod
 CC STREET: 2190 Commons Parkway
 CC CITY: Okemos
 CC STATE: Michigan
 CC COUNTRY: USA
 CC ZIP: 48864
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch, 1.2 MB
 CC FORMATS: IBM PC, MS-DOS 5.0
 CC OPERATING SYSTEM: MS-DOS 5.0
 CC SOFTWARE: PC Write 3.02
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09-132-799-1
 CC FILING DATE: 23-MAY-1996
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 09-0927215
 CC FILING DATE: 23-MAY-1996
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Ian C. McLeod
 CC REGISTRATION NUMBER: 26,931
 CC REFERENCE/DOCKET NUMBER: MSU 4,1-132
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (517) 347-4100
 CC TELEFAX: (517) 347-4100
 CC INFORMATION FOR SEQ ID NO: 14:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 92 amino acids
 CC TYPE: amino acid
 CC STRAIGHTNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: N-terminal fragment
 CC ORIGINAL SOURCE:
 CC ORGANISM: herpes simplex virus type 1 (HSV1)
 CC FEATURE:
 CC NAME, VALUE, LOCATION, 1-16-92
 CC OTHER INFORMATION: Peptide homologous to the US2 gene
 CC OTHER INFORMATION: Peptide of 92 amino acids

SO SEQUENCE 92 AA, 10121 MW, 43655 GN.
 Query Match 100.0%, Score 31, DB 2, Length 92,
 Best Local Similarity 100.0%, Pred No 4 64e-02,
 Matches 4: Conservative 1 Mismatches 0 Indels 0 Gaps 0

Db 50 VWRP 53
 1111
 QY 1 VWRP 4

RESULT 8
 ID US-08-552-815A-14 STANDARD: PRT: 116 AA.
 XX AC xxxxxx
 XX DT
 XX DT

XX Sequence 14, Application US/0852816A
 CC Patent No. 5872215
 CC GENERAL INFORMATION:
 CC APPLICANT: Leand F. Vellard, Peter Brimovskis,
 CC TITLE OF INVENTION: MARSH'S DISCOURSE REPRESENTATIONS
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Ian C. McLeod
 CC STREET: 2190 Commons Parkway
 CC CITY: Okemos
 CC STATE: Michigan
 CC COUNTRY: USA
 CC ZIP: 48864

CC APPLICANT: McCaifferty, JG
 CC TITLE OF INVENTION: MARSH'S DISCOURSE REPRESENTATIONS
 CC NUMBER OF SEQUENCES: 53
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Mark A. McLeod, 233 South Wacker Drive
 CC STREET: 4500 South Wacker, 233 South Wacker Drive
 CC CITY: Chicago
 CC STATE: Illinois
 CC COUNTRY: United States of America
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: MS-DOS/MS-WIN
 CC SOFTWARE: IBM PC compatible
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09-132-799-1
 CC FILING DATE: 23-MAY-1996
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 09-0927215
 CC FILING DATE: 23-MAY-1996
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: David W. Clough
 CC REGISTRATION NUMBER: 36,107
 CC REFERENCE/DOCKET NUMBER: 2411/33308
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (312) 424-1300
 CC INFORMATION FOR SEQ ID NO: 14:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 116 amino acids
 CC TYPE: amino acid
 CC STRAIGHTNESS: linear
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: linear
 CC FRAGMENT TYPE: N-terminal fragment
 CC ORIGINAL SOURCE:
 CC ORGANISM: herpes simplex virus type 1 (HSV1)
 CC FEATURE:
 CC NAME, VALUE, LOCATION, 1-16-92
 CC OTHER INFORMATION: Peptide homologous to the US2 gene
 CC OTHER INFORMATION: Peptide of 116 amino acids

SO SEQUENCE 116 AA, 12860 MW, 59765 GN.
 Query Match 100.0%, Score 31, DB 2, Length 116,
 Best Local Similarity 100.0%, Pred No 4 64e-02,
 Matches 4: Conservative 0 Mismatches 0 Indels 0 Gaps 0

Db 11 VWRP 14
 1111
 QY 1 VWRP 4

RESULT 9
 ID US-08-828-832-3 STANDARD: PRT: 158 AA.
 XX AC xxxxxx
 XX DT
 XX DT

XX Sequence 14, Application US/0852816A
 CC Patent No. 5872215
 CC GENERAL INFORMATION:
 CC APPLICANT: Leand F. Vellard, Peter Brimovskis,
 CC TITLE OF INVENTION: MARSH'S DISCOURSE REPRESENTATIONS
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Ian C. McLeod
 CC STREET: 2190 Commons Parkway
 CC CITY: Okemos
 CC STATE: Michigan
 CC COUNTRY: USA
 CC ZIP: 48864


```

XX Sequence 1 Application US/08865336
CC Patent No. 5958725
CC GENERAL INFORMATION:
CC APPLICANT: Bandman, Olga
CC APPLICANT: Guedler, Karl J.
CC APPLICANT: Shah, Purvi
CC APPLICANT: Pettithory, Joanne R.
CC APPLICANT: Corley, Neil C.
CC TITLE OF INVENTION: NEW HUMAN Dp1 HOMOLOG
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Insyle Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: PASTSEQ for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US-08865336
CC FILING DATE: Herewith
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC AGENCY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/BACKUP NUMBER: PF-0313 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 184 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: STINING102
CC CLONE: 236773
SQ SEQUENCE 184 AA: 20733 MW: 175484 CN:

Query Match 100.0% Score 100.0% length 184
Best Local Similarity 100.0% Pred No. 4.64e+02;
Matches 4, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Db 139 VWRP 142
QY 1 VWRP 4

```

Search completed Fri Feb 25 14:06:44 2000
Job time: 44 secs.

 RELEASE

release 3.1A John P. Collins, Biocomputing Research Unit,
 Copyright (C) 1994-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Release 3.1A John P. Collins, Biocomputing Research Unit,
 Copyright (C) 1994-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Release 3.1A John P. Collins, Biocomputing Research Unit,
 Copyright (C) 1994-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Release 3.1A John P. Collins, Biocomputing Research Unit,
 Copyright (C) 1994-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Release 3.1A John P. Collins, Biocomputing Research Unit,
 Copyright (C) 1994-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Seq. N	Source	Chain	Length	FR ID	Description	Pred. No.
1	1	1	16	PH1778	r cell receptor alpha	2536-02
2	1	2	20	PH1793	r cell receptor beta	2536-02
3	1	3	24	A2732	nicotinic acetylcholine	2536-02
4	1	4	24	A2732	nicotinic acetylcholine	2536-02
5	1	5	66	S4737	hypothetical protein	2536-02
6	1	6	69	J64545	hypothetical protein	2536-02
7	1	7	72	A41889	positive regulator of	2536-02
8	1	8	74	G31174	probable transmembrane	2536-02
9	1	9	82	S75451	hypothetical protein	2536-02
10	1	10	87	G70782	hypothetical protein	2536-02
11	1	11	96	H69042	ribosomal protein L21	2536-02
12	1	12	98	S29228	lg heavy chain V red1	2536-02
13	1	13	100	S16354	hypothetical protein	2536-02
14	1	14	100	S25563	hypothetical protein	2536-02
15	1	15	101	S25563	hypothetical protein	2536-02
16	1	16	101	S25563	hypothetical protein	2536-02
17	1	17	106	G64111	probable transmembrane	2536-02
18	1	18	112	PH1654	lg heavy chain V red1	2536-02
19	1	19	122	F66010	hypothetical protein	2536-02
20	1	20	122	F66009	hypothetical protein	2536-02
21	1	21	123	A31172	trans protein Escher	2536-02
22	1	22	126	S41327	gene M2 protein - Msc	2536-02
23	1	23	142	I34143	aspartic protein precursor	2536-02
24	1	24	144	G75657	hypothetical protein	2536-02

24	31	100.0	146	1	HVWSB1	in heavy chain protein	2536-02
25	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
26	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
27	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
28	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
29	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
30	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
31	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
32	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
33	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
34	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
35	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
36	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
37	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
38	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
39	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
40	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
41	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
42	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
43	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
44	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
45	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
46	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
47	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
48	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
49	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
50	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
51	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
52	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
53	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
54	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
55	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
56	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
57	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
58	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
59	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
60	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
61	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
62	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
63	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
64	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
65	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
66	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
67	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
68	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
69	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
70	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
71	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
72	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
73	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
74	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
75	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
76	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
77	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
78	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
79	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
80	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
81	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
82	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
83	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
84	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
85	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
86	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
87	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
88	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
89	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
90	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
91	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
92	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
93	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
94	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
95	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02
96	31	100.0	146	2	HVWSB1	in heavy chain protein	2536-02

 Release 3.1A John P. Collins, Biocomputing Research Unit,
 Copyright (C) 1994-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

[illegible]

[illegible]

SUMMARY #length 24 #checksum 3911

Query Match 100.0% Score 31 DB 2 Length 24
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 18 VWRP 21
 ||||
 QY 1 VWRP 4

RESULT 4 #type fragment
 ENTRY
 TITLE Nicotinic acetylcholine receptor alpha chain - chicken
 (fragment)
 ORGANISM #formal_name Gallus gallus #common_name chicken
 #journal 20-Jul-1987 #sequence_revision 21-May-1988 #text_change
 DATE 15-Jun-1996

ACCESSION A25738
 REFERENCE A94055
 #authors Conti-Tronconi R M ; Dunn S M ; Barnard E A ; Dolly J O ; Gill P A ; Fajó N ; Fajó M A ;
 #journal Proc Natl Acad Sci U S A (1986) 83:5004-5012
 #title Brain and muscle nicotinic acetylcholine receptors are
 different but homologous proteins
 #cross-references MIM:257384
 #accession A25738

COMMENT #molecule_type Protein
 #residues 124 ##label CON
 CLASSIFICATION The source of this protein was muscle.
 KEYWORDS #superfamily acetylcholine receptor
 glycoprotein ion channel neurotransmitter receptor;
 postsynaptic membrane; transmembrane protein
 #length 24 #checksum 3483

Query Match 100.0% Score 31 DB 2 Length 24
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 17 VWRP 20
 ||||
 QY 1 VWRP 4

RESULT 1 #type complete
 ENTRY
 TITLE Hypothetical protein Acm16 - Anopheles stephensi
 ORGANISM #formal_name Anopheles stephensi
 #journal 16-Mar-1999 #sequence_revision 17-Mar-1999 #text_change
 DATE 17-Mar-1999

ACCESSIONS S45377
 REFERENCE S45377
 #authors Samant S A ; Vilekar N R ; Lole K S ; Deobagkar D D ;
 Deobagkar D N ;
 #journal Indian J Biochem Biophys (1997) 34:423-433
 #title Characterization of Mdr repeat DNA sequence of Anopheles
 stephensi
 #cross-references MIM:609011
 #accession S45377

COMMENT #molecule_type DNA
 #residues 156 ##label SAM
 #cross-references DBL:22758 MIM:609011 17-Mar-1997
 #experimental_source strain fiston
 #length 56 #molecular_weight 7950 #checksum 3911

Query Match 100.0% Score 31 DB 2 Length 56
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 1 VWRP 4
 ||||
 QY 1 VWRP 4

SUMMARY 6

Query Match 100.0% Score 31 DB 2 Length 69
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 35 VWRP 38
 ||||
 QY 1 VWRP 4

RESULT 6 #type complete
 ENTRY
 TITLE Hypothetical 7 k protein - Barophilic bacterium
 ORGANISM #formal_name Barophilic bacterium
 #journal 29-Feb-1996 #sequence_revision 19-Apr-1996 #text_change
 DATE 31-Oct-1997

ACCESSIONS J04545
 REFERENCE J04545
 #authors Kato C ; Smorawinska M ; Sato T ; Horikoshi K ;
 Riosci Biotechnol Biochem (1996) 60:166-168
 #journal Analysis of a pressure-regulated operon from the barophilic
 bacterium strain DB6705.
 #cross-references MIM:267223
 #accession J04545

COMMENT #molecule_type DNA
 #residues 169 ##label KAT
 #cross-references DB:267223 MIM:267223 19-Apr-1996 19-Apr-1996 19-Apr-1996
 #experimental_source DB6705
 #length 69 #checksum 1843

Query Match 100.0% Score 31 DB 2 Length 69
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 35 VWRP 38
 ||||
 QY 1 VWRP 4

SUMMARY 7

Query Match 100.0% Score 31 DB 2 Length 69
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 35 VWRP 38
 ||||
 QY 1 VWRP 4

RESULT 7 #type complete
 ENTRY
 TITLE Positive regulator of phase P2 late gene transcription -
 Escherichia coli
 ORGANISM #formal_name Escherichia coli
 #journal 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
 DATE 17-Mar-1999

ACCESSIONS A41889
 REFERENCE A41889
 #authors Barreiro V ; Haggard-Ljungquist E ;
 J Bacteriol (1992) 174:4086-4093
 #journal Attachment sites for bacteriophage P2 on the Escherichia coli
 chromosome: DNA sequences localization on the physical
 map, and detection of a P2-like remnant in Escherichia coli
 K-12 derivatives.
 #cross-references MIM:257384

Query Match 100.0% Score 31 DB 2 Length 69
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 35 VWRP 38
 ||||
 QY 1 VWRP 4

RESULT 7 #type complete
 ENTRY
 TITLE Positive regulator of phase P2 late gene transcription -
 Escherichia coli
 ORGANISM #formal_name Escherichia coli
 #journal 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
 DATE 17-Mar-1999

ACCESSIONS A41889
 REFERENCE A41889
 #authors Barreiro V ; Haggard-Ljungquist E ;
 J Bacteriol (1992) 174:4086-4093
 #journal Attachment sites for bacteriophage P2 on the Escherichia coli
 chromosome: DNA sequences localization on the physical
 map, and detection of a P2-like remnant in Escherichia coli
 K-12 derivatives.
 #cross-references MIM:257384

Query Match 100.0% Score 31 DB 2 Length 69
 Best Local Similarity 100.0% Pred No 2 Sep-92
 Matches 4: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 35 VWRP 38
 ||||
 QY 1 VWRP 4

[illegible]

```

Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

DE      75 VWRP 78
      ||||
QY      1 VWRP 4

RESULT 11
ENTRY
TITLE  ribosomal protein L21 - Methanobacterium thermoautotrophicum
      (strain Delta H)
ORGANISM  #formal_name Methanobacterium thermoautotrophicum
DATE  05-Dec-1993 #sequence_revision 05 Feb-1997 #text_change
      05-Jun-1998
ACCESSIONS  H59042
REFERENCE  A59000
#authors  Smith, D.F.; Doucette-Stamm, L.A.; Deloughery, G.; Lee, H.;
      Dubois, J.; Alarid, J.; Bashir-Zadeh, S.; Plakke, P.;
      Cook, P.; Gilbert, W.; Harrison, D.; Hoang, I.; Keagle, P.;
      Lumm, W.; Pothier, R.; Qiu, D.; Spadafora, P.; Vicari, P.;
      Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwanli, N.; Caruso,
      A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;
      McLaughlin, J.; Shirai, S.; Goyal, A.; Piotrowski, S.;
      Church, G.W.; Cantale, G.T.; Naei, C.; Rice, F.; Neill,
      J.; Reeve, J.N.
#journal  J. Bacteriol. (1997) 179:7135-7155
#title  Complete genome sequence of Methanobacterium
      thermoautotrophicum Delta H: functional analysis and
      comparative genomics.
#cross-references  MITH:98037514
#accession  H59042
#status  Preliminary; nucleic acid sequence not shown;
      translation not shown
#molecule_type  DNA
#residues  1-96 #label MTH
#cross-references  EF AB019906, EF AB019907, EF 9822424, EF 9822423
#experimental_source  strain Delta H
GENETICS
#gene  MTH1323
CLASSIFICATION  #superfamily Halcaerulea ribosomal protein HL31
SUMMARY  #length 96 #molecular_weight 11116 #checksum 6593

Query Match      100.0%  Score 31:  DB 2:  Length 96:
Best Local Similarity 100.0%  Pred No 2: 58e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

DE      86 VWRP 84
      ||
QY      1 VWRP 4

RESULT 12
ENTRY
TITLE  1d heavy chain V region (DP-29) - human (fragment)
ORGANISM  #formal_name Homo sapiens #common_name man
DATE  22-Nov-1993 #sequence_revision 10 Nov-1995 #text_change
      17-Mar-1999
ACCESSIONS  S26928
REFERENCE  S26985
#authors  Tomlinson, I. M.; Walter, G. S.; Marks, J. P.; Hleawlyn, M. R.;
      Winter, G.
#journal  J. Mol. Biol. (1997) 277:77-90
#title  The repertoire of human germline V(H) sequences reveals about
      fifty groups of V(H) segments with different hypervariable
      loops.
#cross-references  MIM:601117
#accession  S26928
#status  Preliminary
#molecule_type  DNA
#residues  1-98 #label IOM
#cross-references  EMBL 912944, NIG 930487, F01932889
CLASSIFICATION  #superfamily immunoglobulin V region: immunoglobulin homology
KEYWORDS  heterotrimer: immunoglobulin

```

```

FEATURE
15-98      #domain immunoglobulin homology #label IMM
SUMMARY  #length 98 #checksum 2406

Query Match      100.0%  Score 31:  DB 2:  Length 98:
Best Local Similarity 100.0%  Pred No 2: 53e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

DE      11 VWRP 14
      ||||
QY      1 VWRP 4

RESULT 13
ENTRY
TITLE  Hypothetical protein 2 - phage phi-C31
ORGANISM  #formal_name Phage phi-C31
DATE  20 Feb-1995 #sequence_revision 20-Feb-1995 #text_change
      04-Sep-1997
ACCESSIONS  S16054
REFERENCE  S16053
#authors  Rausch, H.
#journal  J. Virol. (1995) 69:14810
#title  EMVAT Data Library December 1990
#accession  S16054
#status  Preliminary
#molecule_type  DNA
#residues  1-100 #label RAU
#cross-references  EMBL X60846, NIG 914810, F009579069
GENETICS
#start_codon  GTG
SUMMARY  #length 100 #molecular_weight 11548 #checksum 22

Query Match      100.0%  Score 31:  DB 2:  Length 100:
Best Local Similarity 100.0%  Pred No 2: 53e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

DE      30 VWRP 33
      ||||
QY      1 VWRP 4

RESULT 14
ENTRY
TITLE  Hypothetical protein 3 - phage phi-C31
ORGANISM  #formal_name Phage phi-C31
DATE  22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
      10-Nov-1995
ACCESSIONS  S22567
REFERENCE  S22566
#authors  Rausch, H.; Lehmann, M.
#journal  Nucleic Acids Res. (1991) 19:5187-5189
#title  Structural analysis of the actinophage phiC31 attachment
      site.
#cross-references  MIM:62000093
#accession  S22567
#status  Preliminary; nucleic acid sequence not shown;
      translation not shown
#molecule_type  DNA
#residues  1-100 #label RAU
#cross-references  EMBL X60846
#note  the nucleotide sequence was submitted to the EMBL Data
      Library December 1990
SUMMARY  #length 100 #molecular_weight 11516 #checksum 41

Query Match      100.0%  Score 31:  DB 2:  Length 100:
Best Local Similarity 100.0%  Pred No 2: 53e-02:
Matches      4:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

DE      30 VWRP 33
      ||||
QY      1 VWRP 4

```




 WISRELI

 (18)

Release 6.0A John F. Collins, BioComputing Research Unit.

Copyright (c) 1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

WISRELI is a protein database search, using Smith-Waterman algorithm

Run on Fri Feb 25 14:06:42 2000; MasPar time 2.69 Seconds
 56.566 Million cell updates/sec

Database used is protein.

Query: 000-06-13-749-1
 (14) from RS09142799.pep
 Length: 31
 AVSF 4

Sequence: IAM 15
 Gap 15

Score: 175.77 seqs, 2426243 residues

Query description: Mus musculus

Query title: Distinct first 1000 summaries

Query type: Extensive
 SwissProt

Statistics: Mean 18.116; Variance 13.185; scale 1.374

Note: No. in the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

Seq. No.	Score	Match	Length	FR	ID	Description	Pred. No.
1	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
2	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
3	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
4	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
5	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
6	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
7	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
8	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
9	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
10	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
11	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
12	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
13	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
14	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
15	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
16	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
17	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
18	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
19	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02
20	175	24	31	1	1	ACTIVATOR PROTEIN RECEPTOR	1,036-02

214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384	1385	1386	1387	1388	1389	1390	1391	1392	1393	1394	1395	1396	1397	1398	1399	1400	1401	1402	1403	1404	1405	1406	1407	1408	1409	1410	1411	1412	1413	1414	1415	1416	1417	1418	1419	1420	1421	1422	1423	1424	1425	1426	1427	1428	1429	1430	1431	1432	1433	1434	1435	1436	1437	1438	1439	1440	1441	1442	1443	1444	1445	1446	1447	1448	1449	1450	1451	1452	1453	1454	1455	1456	1457	1458	1459	1460	1461	1462	1463	1464	1465	1466	1467	1468	1469	1470	1471	1472	1473	1474	1475	1476	1477	1478	1479	1480	1481	1482	1483	1484	1485	1486	1487	1488	1489	1490	1491	1492	1493	1494	1495	1496	1497	1498	1499	1500	1501	1502	1503	1504	1505	1506	1507	1508	1509	1510	1511	1512	1513	1514	1515	1516	1517	1518	1519	1520	1521	1522	1523	1524	1525	1526	1527	1528	1529	1530	1531	1532	1533	1534	1535	1536	1537	1538	1539	1540	1541	1542	1543	1544	1545	1546	1547	1548	1549	1550	1551	1552	1553	1554	1555	1556	1557	1558	1559	1560	1561	1562	1563	1564	1565	1566	1567	1568	1569	1570	1571	1572	1573	1574	1575	1576	1577	1578	1579	1580	1581	1582	1583	1584	1585	1586	1587	1588	1589	1590	1591	1592	1593	1594	1595	1596	1597	1598	1599	1600	1601	1602	1603	1604	1605	1606	1607	1608	1609	1610	1611	1612	1613	1614	1615	1616	1617	1618	1619	1620	1621	1622	1623	1624	1625	1626	1627	1628	1629	1630	1631	1632	1633	1634	1635	1636	1637	1638	1639	1640	1641	1642	1643	1644	1645	1646	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659</
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

[illegible]

RA SODI J.D., SHIRLEY R., LIU L.-I., GLENN A., KELLEY J.M.,
 RA WEGMAN C.F., PHILLIPS D.A., PHILLIPS D., HELLMAN E., ELIEN H.L.,
 RA CARRERA V.E., HARRA M., WINTER E., FANTER D.W., BRANTON S.E.,
 RA FINE L.D., FRECHMAN S.L., FUERNBERG S.L., GOSCHAGEN R.S.M.,
 RA GRIMM C.L., MEDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
 RA VENTER J.C.,
 RA "RNA POLYMERASE TRANSCRIPTION AND ASSEMBLY OF HOST-CELLS
 RA INFLUENZA B",
 RA SCIENCE 257:450-451 (1995),
 RA 41. SIMILARITY: BELONGS TO THE SUI1 FAMILY.
 RA
 RA This Swiss-Prot entry is copyright. It is produced through a collaboration
 RA between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 RA the European Bioinformatics Institute. There are no restrictions on its
 RA use by non-profit institutions as long as its content is in no way
 RA altered. If this agreement is not removed, "SUI1" by and for commercial
 RA entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 RA or send an email to license@sib-sib.ch)
 RA
 RA EMBL: U32802; G1574156; -
 RA TIGR: H11225; -
 RA PROSITE: PS01118; SUI1; 1;
 RA PRAM: PF01253; SUI1; 1;
 KW HYPOTHEETICAL PROTEIN: PROTEIN BIOSYNTHESIS, TRANSLATION REGULATION
 SQ SEQUENCE 106 AA; 12226 MW; 855-1932 CRC32;
 Query Match 100.0%; Score 31; E-1 Length 106.
 Best Local Similarity 100.0%; Pred. No. 103e-02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 22 VWRP 24
 QY 1 VWRP 4
 RESULT 8
 ID YFDO_ECOLI STANDARD; FR1; 122 AA;
 AC P17909;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DI 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHEETICAL 14.4 KD PROTEIN IN INT3-DSDC INTERGENIC REGION.
 GN YFDO.
 OS ESCHERICHIA COLI.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 OC ESCHERICHIA.
 RN [1]
 PP SEQUENCE FROM N.A. AND SEQUENCE OF 1-5 AND 118-122.
 PX MEDLINE: 8930605
 RA ZIEGLIN G., FUERST J.P., LANKA E.;
 RA "17.5 kDa protein of plasmid RP4 binds to a 19-base pair invert sequence
 RA repetition within the transfer origin";
 RA J. BIOL. CHEM. 264:11989-11994 (1989).
 RN [2]
 PP SEQUENCE FROM N.A.
 AC STRAIN=H101;
 EX MEDLINE: 92190548;
 RA ZIEGLIN G., FANSEGRAD W., STRACK E., PALZER D., REGER M., FRUT V.;
 RA LANKA E.;
 RA "Nucleotide sequence and organization of genes flanking the transfer
 RA origin of promiscuous plasmid RP4";
 RA DNA SEQ. 1:303-327 (1981).
 RN [3]
 PP SEQUENCE OF 1-7 FROM N.A.
 AC MEDLINE: 89184510;
 RA LANKA E., PURPSTE J.P.;
 RA "Cloning, transfer, and plasmid transfer of plasmids: initiation of
 RA plasmid-encoded products with the transfer origin";
 RA EMBO J. 13:1791-1795 (1994).
 RA
 RA FUNCTION: TRANSFER OF PLASMID RP4 DURING BACTERIAL CONJUGATION
 RA REQUIRES THE PLASMID-ENCODED TRAP PROTEIN, WHICH BINDS TO A 19-
 RA BASE PAIR INVERT SEQUENCE REPEITION WITHIN THE TRANSFER ORIGIN.
 RA TRAP PROTEIN IS BOUND TO ONLY ONE SIDE OF THE DNA HELIX. THIS
 RA NUCLEOPROTEIN STRUCTURE IS THE INITIAL COMPLEX IN THE PATHWAY TO
 RA ASSEMBLY A FUNCTIONAL REPLICASOME.
 RA
 RA SURUNIT: MONOMER.
 RA
 RA This Swiss-Prot entry is copyright. It is produced through a collaboration
 RA between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 RA the European Bioinformatics Institute. There are no restrictions on its
 RA use by non-profit institutions as long as its content is in no way
 RA altered. If this agreement is not removed, "SUI1" by and for commercial
 RA entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 RA or send an email to license@sib-sib.ch)
 RA
 RA EMBL: U34342; G45934; -
 RA EMBL: X54459; G45785; -

RA use by non-profit institutions as long as its content is in no way
 RA altered. If this agreement is not removed, "SUI1" by and for commercial
 RA entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 RA or send an email to license@sib-sib.ch)
 RA
 RA EMBL: AF000324; G1789700; -
 RA FMRK; FMRK; 3179773; ALT_INIT;
 RA ECGENEL; EGI4137; YFDO;
 RA HYPOTHEETICAL PROTEIN;
 SQ SEQUENCE 122 AA; 14431 MW; 15793656 CRC32;
 Query Match 100.0%; Score 31; E-1 Length 122;
 Best Local Similarity 100.0%; Pred. No. 103e-02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 22 VWRP 24
 QY 1 VWRP 4
 RESULT 9
 ID YFDO_ECOLI STANDARD; FR1; 122 AA;
 AC P17909;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DI 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE TRAP PROTEIN (RELAX-SOME PROTEIN).
 GN TRAP;
 OS ESCHERICHIA COLI.
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
 OC ESCHERICHIA.
 RN [1]
 PP SEQUENCE FROM N.A. AND SEQUENCE OF 1-5 AND 118-122.
 PX MEDLINE: 8930605
 RA ZIEGLIN G., FUERST J.P., LANKA E.;
 RA "17.5 kDa protein of plasmid RP4 binds to a 19-base pair invert sequence
 RA repetition within the transfer origin";
 RA J. BIOL. CHEM. 264:11989-11994 (1989).
 RN [2]
 PP SEQUENCE FROM N.A.
 AC STRAIN=H101;
 EX MEDLINE: 92190548;
 RA ZIEGLIN G., FANSEGRAD W., STRACK E., PALZER D., REGER M., FRUT V.;
 RA LANKA E.;
 RA "Nucleotide sequence and organization of genes flanking the transfer
 RA origin of promiscuous plasmid RP4";
 RA DNA SEQ. 1:303-327 (1981).
 RN [3]
 PP SEQUENCE OF 1-7 FROM N.A.
 AC MEDLINE: 89184510;
 RA LANKA E., PURPSTE J.P.;
 RA "Cloning, transfer, and plasmid transfer of plasmids: initiation of
 RA plasmid-encoded products with the transfer origin";
 RA EMBO J. 13:1791-1795 (1994).
 RA
 RA FUNCTION: TRANSFER OF PLASMID RP4 DURING BACTERIAL CONJUGATION
 RA REQUIRES THE PLASMID-ENCODED TRAP PROTEIN, WHICH BINDS TO A 19-
 RA BASE PAIR INVERT SEQUENCE REPEITION WITHIN THE TRANSFER ORIGIN.
 RA TRAP PROTEIN IS BOUND TO ONLY ONE SIDE OF THE DNA HELIX. THIS
 RA NUCLEOPROTEIN STRUCTURE IS THE INITIAL COMPLEX IN THE PATHWAY TO
 RA ASSEMBLY A FUNCTIONAL REPLICASOME.
 RA
 RA SURUNIT: MONOMER.
 RA
 RA This Swiss-Prot entry is copyright. It is produced through a collaboration
 RA between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 RA the European Bioinformatics Institute. There are no restrictions on its
 RA use by non-profit institutions as long as its content is in no way
 RA altered. If this agreement is not removed, "SUI1" by and for commercial
 RA entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 RA or send an email to license@sib-sib.ch)
 RA
 RA EMBL: U34342; G45934; -
 RA EMBL: X54459; G45785; -

```

14 EMBL: M274.1 (F55174) ALL_INT.
15 EMBL: A34771 (A34771)
16 EMBL: S43113 (S43113)
17 EMBL: S43113 (S43113)
18 EMBL: S43113 (S43113)
19 EMBL: S43113 (S43113)
20 EMBL: S43113 (S43113)
21 EMBL: S43113 (S43113)
22 EMBL: S43113 (S43113)
23 EMBL: S43113 (S43113)
24 EMBL: S43113 (S43113)
25 EMBL: S43113 (S43113)
26 EMBL: S43113 (S43113)
27 EMBL: S43113 (S43113)
28 EMBL: S43113 (S43113)
29 EMBL: S43113 (S43113)
30 EMBL: S43113 (S43113)
31 EMBL: S43113 (S43113)
32 EMBL: S43113 (S43113)
33 EMBL: S43113 (S43113)
34 EMBL: S43113 (S43113)
35 EMBL: S43113 (S43113)
36 EMBL: S43113 (S43113)
37 EMBL: S43113 (S43113)
38 EMBL: S43113 (S43113)
39 EMBL: S43113 (S43113)
40 EMBL: S43113 (S43113)
41 EMBL: S43113 (S43113)
42 EMBL: S43113 (S43113)
43 EMBL: S43113 (S43113)
44 EMBL: S43113 (S43113)
45 EMBL: S43113 (S43113)
46 EMBL: S43113 (S43113)
47 EMBL: S43113 (S43113)
48 EMBL: S43113 (S43113)
49 EMBL: S43113 (S43113)
50 EMBL: S43113 (S43113)
51 EMBL: S43113 (S43113)
52 EMBL: S43113 (S43113)
53 EMBL: S43113 (S43113)
54 EMBL: S43113 (S43113)
55 EMBL: S43113 (S43113)
56 EMBL: S43113 (S43113)
57 EMBL: S43113 (S43113)
58 EMBL: S43113 (S43113)
59 EMBL: S43113 (S43113)
60 EMBL: S43113 (S43113)
61 EMBL: S43113 (S43113)
62 EMBL: S43113 (S43113)
63 EMBL: S43113 (S43113)
64 EMBL: S43113 (S43113)
65 EMBL: S43113 (S43113)
66 EMBL: S43113 (S43113)
67 EMBL: S43113 (S43113)
68 EMBL: S43113 (S43113)
69 EMBL: S43113 (S43113)
70 EMBL: S43113 (S43113)
71 EMBL: S43113 (S43113)
72 EMBL: S43113 (S43113)
73 EMBL: S43113 (S43113)
74 EMBL: S43113 (S43113)
75 EMBL: S43113 (S43113)
76 EMBL: S43113 (S43113)
77 EMBL: S43113 (S43113)
78 EMBL: S43113 (S43113)
79 EMBL: S43113 (S43113)
80 EMBL: S43113 (S43113)
81 EMBL: S43113 (S43113)
82 EMBL: S43113 (S43113)
83 EMBL: S43113 (S43113)
84 EMBL: S43113 (S43113)
85 EMBL: S43113 (S43113)
86 EMBL: S43113 (S43113)
87 EMBL: S43113 (S43113)
88 EMBL: S43113 (S43113)
89 EMBL: S43113 (S43113)
90 EMBL: S43113 (S43113)
91 EMBL: S43113 (S43113)
92 EMBL: S43113 (S43113)
93 EMBL: S43113 (S43113)
94 EMBL: S43113 (S43113)
95 EMBL: S43113 (S43113)
96 EMBL: S43113 (S43113)
97 EMBL: S43113 (S43113)
98 EMBL: S43113 (S43113)
99 EMBL: S43113 (S43113)
100 EMBL: S43113 (S43113)

```


10 between the Swiss Institute of Bioinformatics and the EMBL outstation.
 11 The EMBL outstation is not a full-fledged institution. There are no restrictions on its
 12 use for non-profit institutions as long as its content is in no way
 13 related to this statement is not removed. Usage by and for commercial
 14 entities is prohibited. Please inform the EMBL outstation (see above) if you
 15 are not an email address (otherwise list it).

16 -----
 17 DE EMBL-OUTSTATION (1992-02-25)
 18 DE EMBL-OUTSTATION (1992-02-25)
 19 DE EMBL-OUTSTATION (1992-02-25)

20 DE EMBL-OUTSTATION (1992-02-25) 22RRR70C CRC32:

21 From: Martin J. (mailto:Martin.J@embl.ch) Date: 1992-02-25
 22 Subject: Similarity (0.000) Pred. No. 1.000-02:
 23 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

24 11 44 VWRP 47
 25 11 1 VWRP 4

26 Sent: 1992-02-25 14:06:48 Fri Feb 25 13:07:00 2000
 27 Subject: 47 spars.



Revised: J. A. John F. Collins, BioComputing Research Unit,
Department of Chemistry, University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Memo 4.11
mooC++ - proof database search, using Smith-Waterman algorithm
Fri Feb 26 1987 Tue Mar 2 1987
Master time 4.23 Seconds
51,645 Million words processed

1. *Author*: [Name]
 2. *Title*: [Title]
 3. *Journal*: [Journal]
 4. *Volume*: [Volume]
 5. *Issue*: [Issue]
 6. *Page*: [Page]
 7. *Year*: [Year]
 8. *DOI*: [DOI]
 9. *URL*: [URL]
 10. *Accessed*: [Accessed]

604: 29-41 11,766 sq. ft., 545,874 residues
 605: 1-3 associated Maximum M.P. 309
 first two first 1000 summaries

sp1000000

20.0	13.1	8.3	Mean	17.746	Variance	13.372	scale	3.327
------	------	-----	------	--------	----------	--------	-------	-------

ferred. Now, is the number of results predicted by chance to have a value greater than or equal to the score of the result being printed, and is the probability of the total score distribution.

SUMMARY

Accession No.	Entry	Gene	Ref.	Length	MB	IC	Protein	Prod. No.
1	41	10010	46	2	066740		PRESSURE-REGULATED GRI	2,050-02
2	41	10010	46	2	066768		PRESSURE-REGULATED GRI	2,050-02
3	41	10010	47	5	026196		IMMUNOGENIC ANTIGEN	2,050-02
4	41	10010	47	5	026255		UNKNOWN PROTEIN (FRAGMENT)	2,050-02
5	41	10010	48	5	017436		HYPOTHEICAL 7.8 KD PR	2,050-02
6	41	10010	49	2	076736		PRESSURE REGULATED GRI	2,050-02
7	41	10010	50	2	074368		RICE, CYD5, CYD6 AND P	2,050-02
8	41	10010	74	1	074091		74AA LONG HYPOTHEICAL	2,050-02
9	41	10010	75	2	005168		HYPOTHEICAL 8.4 KD PR	2,050-02
10	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
11	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
12	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
13	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
14	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
15	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
16	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
17	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
18	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
19	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
20	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
21	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
22	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
23	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
24	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
25	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
26	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
27	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
28	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
29	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
30	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
31	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
32	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
33	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
34	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
35	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
36	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
37	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
38	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
39	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
40	41	10010	81	14	046602		STRAIN KIEGIC12A (25H)	2,050-02
41	41	10010	81	14	046602			


```

PT STAIN "HEPATITIS (TASHKENT)" DBF-1 (FRAGMENT).
OS HEPATITIS E VIRUS (HEV).
OC VIRUSES: SSRNA POSITIVE STRAND VIRUSES, NO DNA STAGE. CALICIVIRIDAE.
OC CALICIVIRUS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=UZBEKISTAN (TASHKENT);
RC CHATIRIJE R., TSAJEV S., PILOT J., COURSAJET P., EMERSON S.,
RA PURCELL R.;
RF 1 MED VIRUS 0-0-0(1997).
RL EMBL AF010426; G322319; -.
RS 1
RT NON-IFR 1
FT NON-IFR 81
FS SEQUENCE 81 AA: 2588 MW: 507797 257324
Query Match 100.00 Score 31. DB 14; Length 81;
Pos. 1-81 Similarity 100.00 Pval No 2.05e-02;
Matches 4; Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 2 VWRP 5
||||
CY 1 VWRP 4

RESULT 12
ID C36600 PRELIMINARY: PRT: 81 AA.
AC G36600:
DI 01-JAN-1998 (TREMREL. 05, CREATED)
DI 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DI 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
RE STAIN "HEPATITIS (TASHKENT)" DBF-1 (FRAGMENT).
OC HEPATITIS E VIRUS (HEV).
OS VIRUSES: SSRNA POSITIVE STRAND VIRUSES, NO DNA STAGE. CALICIVIRIDAE.
OC CALICIVIRUS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MOROCCO/F12;
RC CHATIRIJE R., TSAJEV S., PILOT J., COURSAJET P., EMERSON S.,
RA PURCELL R.;
RF 1 MED VIRUS 0-0-0(1997).
RL EMBL AF010423; G322313; -.
RS 1
RT NON-IFR 1
FT NON-IFR 81
FS SEQUENCE 81 AA: 2574 MW: 507797 257324
Query Match 100.00 Score 31. DB 14; Length 81;
Pos. 1-81 Similarity 100.00 Pval No 2.05e-02;
Matches 4; Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 2 VWRP 5
||||
QY 1 VWRP 4

RESULT 13
ID C36601 PRELIMINARY: PRT: 81 AA.
AC G36601:
DI 01-JAN-1998 (TREMREL. 05, CREATED)
DI 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DI 01-JAN-1998 (TREMREL. 05, LAST ANNOTATION UPDATE)
RE STAIN "HEPATITIS (TASHKENT)" DBF-1 (FRAGMENT).
OC HEPATITIS E VIRUS (HEV).
OS VIRUSES: SSRNA POSITIVE STRAND VIRUSES, NO DNA STAGE. CALICIVIRIDAE.
OC CALICIVIRUS.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MOROCCO/F23;
RC CHATIRIJE R., TSAJEV S., PILOT J., COURSAJET P., EMERSON S.,
RA PURCELL R.;
RF 1 MED VIRUS 0-0-0(1997).
RL EMBL AF010424; G322315; -.
RS 1
RT NON-IFR 1
FT NON-IFR 81

```

Seq: SEQUENCE 14 AA: 8674 MW: 68A7673 CPG32;
 Query Match 100.0% Score 31; DB 14; Length 81;
 Best Local Similarity 100.0% Pred. No. 2.05e-02;
 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

11 1 VWRP 4
 12 1 VWRP 4

Seq: SEQUENCE 14
 11 1 VWRP 4
 12 1 VWRP 4
 13 1 VWRP 4
 14 1 VWRP 4
 15 1 VWRP 4
 16 1 VWRP 4
 17 1 VWRP 4
 18 1 VWRP 4
 19 1 VWRP 4
 20 1 VWRP 4
 21 1 VWRP 4
 22 1 VWRP 4
 23 1 VWRP 4
 24 1 VWRP 4
 25 1 VWRP 4
 26 1 VWRP 4
 27 1 VWRP 4
 28 1 VWRP 4
 29 1 VWRP 4
 30 1 VWRP 4
 31 1 VWRP 4
 32 1 VWRP 4
 33 1 VWRP 4
 34 1 VWRP 4
 35 1 VWRP 4
 36 1 VWRP 4
 37 1 VWRP 4
 38 1 VWRP 4
 39 1 VWRP 4
 40 1 VWRP 4
 41 1 VWRP 4
 42 1 VWRP 4
 43 1 VWRP 4
 44 1 VWRP 4
 45 1 VWRP 4
 46 1 VWRP 4
 47 1 VWRP 4
 48 1 VWRP 4
 49 1 VWRP 4
 50 1 VWRP 4
 51 1 VWRP 4
 52 1 VWRP 4
 53 1 VWRP 4
 54 1 VWRP 4
 55 1 VWRP 4
 56 1 VWRP 4
 57 1 VWRP 4
 58 1 VWRP 4
 59 1 VWRP 4
 60 1 VWRP 4
 61 1 VWRP 4
 62 1 VWRP 4
 63 1 VWRP 4
 64 1 VWRP 4
 65 1 VWRP 4
 66 1 VWRP 4
 67 1 VWRP 4
 68 1 VWRP 4
 69 1 VWRP 4
 70 1 VWRP 4
 71 1 VWRP 4
 72 1 VWRP 4
 73 1 VWRP 4
 74 1 VWRP 4
 75 1 VWRP 4
 76 1 VWRP 4
 77 1 VWRP 4
 78 1 VWRP 4
 79 1 VWRP 4
 80 1 VWRP 4
 81 1 VWRP 4
 82 1 VWRP 4
 83 1 VWRP 4
 84 1 VWRP 4
 85 1 VWRP 4
 86 1 VWRP 4
 87 1 VWRP 4
 88 1 VWRP 4
 89 1 VWRP 4
 90 1 VWRP 4
 91 1 VWRP 4
 92 1 VWRP 4
 93 1 VWRP 4
 94 1 VWRP 4
 95 1 VWRP 4
 96 1 VWRP 4
 97 1 VWRP 4
 98 1 VWRP 4
 99 1 VWRP 4
 100 1 VWRP 4

Query Match 100.0% Score 31; DB 2; Length 82;
 Best Local Similarity 100.0% Pred. No. 2.05e-02;
 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

11 1 VWRP 4
 12 1 VWRP 4

Seq: SEQUENCE 14
 11 1 VWRP 4
 12 1 VWRP 4
 13 1 VWRP 4
 14 1 VWRP 4
 15 1 VWRP 4
 16 1 VWRP 4
 17 1 VWRP 4
 18 1 VWRP 4
 19 1 VWRP 4
 20 1 VWRP 4
 21 1 VWRP 4
 22 1 VWRP 4
 23 1 VWRP 4
 24 1 VWRP 4
 25 1 VWRP 4
 26 1 VWRP 4
 27 1 VWRP 4
 28 1 VWRP 4
 29 1 VWRP 4
 30 1 VWRP 4
 31 1 VWRP 4
 32 1 VWRP 4
 33 1 VWRP 4
 34 1 VWRP 4
 35 1 VWRP 4
 36 1 VWRP 4
 37 1 VWRP 4
 38 1 VWRP 4
 39 1 VWRP 4
 40 1 VWRP 4
 41 1 VWRP 4
 42 1 VWRP 4
 43 1 VWRP 4
 44 1 VWRP 4
 45 1 VWRP 4
 46 1 VWRP 4
 47 1 VWRP 4
 48 1 VWRP 4
 49 1 VWRP 4
 50 1 VWRP 4
 51 1 VWRP 4
 52 1 VWRP 4
 53 1 VWRP 4
 54 1 VWRP 4
 55 1 VWRP 4
 56 1 VWRP 4
 57 1 VWRP 4
 58 1 VWRP 4
 59 1 VWRP 4
 60 1 VWRP 4
 61 1 VWRP 4
 62 1 VWRP 4
 63 1 VWRP 4
 64 1 VWRP 4
 65 1 VWRP 4
 66 1 VWRP 4
 67 1 VWRP 4
 68 1 VWRP 4
 69 1 VWRP 4
 70 1 VWRP 4
 71 1 VWRP 4
 72 1 VWRP 4
 73 1 VWRP 4
 74 1 VWRP 4
 75 1 VWRP 4
 76 1 VWRP 4
 77 1 VWRP 4
 78 1 VWRP 4
 79 1 VWRP 4
 80 1 VWRP 4
 81 1 VWRP 4
 82 1 VWRP 4
 83 1 VWRP 4
 84 1 VWRP 4
 85 1 VWRP 4
 86 1 VWRP 4
 87 1 VWRP 4
 88 1 VWRP 4
 89 1 VWRP 4
 90 1 VWRP 4
 91 1 VWRP 4
 92 1 VWRP 4
 93 1 VWRP 4
 94 1 VWRP 4
 95 1 VWRP 4
 96 1 VWRP 4
 97 1 VWRP 4
 98 1 VWRP 4
 99 1 VWRP 4
 100 1 VWRP 4

Query Match 100.0% Score 31; DB 2; Length 82;

Best Local Similarity 100.0% Pred. No. 2.05e-02;
 Matches 4; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

DB 35 VWRP 48
 11 1 VWRP 4
 12 1 VWRP 4

Search completed: Fri Feb 25 14:08:45 2000
 Job time : 67 secs.



Release 3.0A John F. Collins, Bioinformatics Research Unit.

Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

Waterman Smith-Waterman algorithm

Score: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

Label: 32.709 Million cell updates/sec

CC Specific uses are described for each of the polypeptides based on
 CC which tissues they are most highly expressed in, and the use of these
 CC products for the diagnosis or treatment of cancer, neurodegenerative
 CC disorders, developmental abnormalities and fetal deficiencies, viral
 CC diseases, tumors, leukemias, diseases of the immune system, autoimmune
 CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
 CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
 CC osteoporosis, prostate diseases, obesity, disorders involving osteoclasts
 CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
 CC and liver, digestive disorders, disorders of the eye, and AIDS. The
 CC polypeptides are also useful for identifying their binding partners.
 CC The present sequence represents a predicted, fragment encoded by a
 CC gene of the invention (see descriptor line for gene number).
 CC Sequence 37 AA:
 CC
 CC Query Match 94.94 Score 37. DR 38. Length 37:
 CC Best Local Similarity 80.84 Prev. No. 2.98e-02
 CC Matches 4. Conservative 1. Mismatches 0. Indels 0. Gaps 0:
 CC
 CC Lb 32 Virpp 36
 CC 1 VVRPP 5
 CC
 CC PPSULT 9
 CC W6065 Standard. Protein: 170 AA.
 CC
 CC DE Polypeptide fragment encoded by gene 190.
 CC Human, secreted protein; fusion protein; gene therapy; protein therapy;
 CC diagnosis, tissue, cancer, tumor, neurodegenerative disorder, leukaemia;
 CC developmental abnormality; fetal deficiency; blood, allergy; renal;
 CC immune system, asthma, lymphocytic disease, brain, hepatitis, lymphoma;
 CC inflammation, ischemic shock; Alzheimer's disease; osteoporosis; AIDS;
 CC osteoporosis; arthritis; testis, lung, thyroiditis; thyroid; digestion;
 CC regulation of the skeletal system; prostate; obesity; osteoclast; thymus;
 CC endocrine metabolism; regulation, malabsorption, gastritis, neoplasm.
 CC
 CC OS Homo sapiens.
 CC NCBI accession: W6065A963-A2.
 CC
 CC 04-JUN-1997: 011422.
 CC 18-DEC-1997: US-070623.
 CC 06-JUN-1997: US-048877.
 CC 06-JUN-1997: US-048881.
 CC 06-JUN-1997: US-048884.
 CC 06-JUN-1997: US-048893.
 CC 06-JUN-1997: US-048896.
 CC 06-JUN-1997: US-048899.
 CC 06-JUN-1997: US-048915.
 CC 06-JUN-1997: US-048949.
 CC 06-JUN-1997: US-048972.
 CC 06-JUN-1997: US-049070.
 CC 06-JUN-1997: US-049375.
 CC 05-SEP-1997: US-057628.
 CC 05-SEP-1997: US-057635.
 CC 05-SEP-1997: US-057644.
 CC 05-SEP-1997: US-057647.
 CC 05-SEP-1997: US-057650.
 CC 05-SEP-1997: US-057661.
 CC 05-SEP-1997: US-057667.
 CC 05-SEP-1997: US-057671.
 CC 05-SEP-1997: US-057770.
 CC 05-SEP-1997: US-057775.
 CC 05-SEP-1997: US-057778.
 CC 06-JUN-1997: US-048875.
 CC 06-JUN-1997: US-048878.
 CC 06-JUN-1997: US-048882.
 CC 06-JUN-1997: US-048885.
 CC 06-JUN-1997: US-048894.
 CC 06-JUN-1997: US-048897.
 CC 06-JUN-1997: US-048900.
 CC
 CC PI Brewer LA, Carter KC, Dillon PJ, Ebner P, Fudress GA,
 CC Fan P, Feng P, Ferlie AM, Fischer CL, Florence C,
 CC Florence V, Greene TM, Hu Y, Kryw H, Laflour DW,
 CC Li Y, Moore PA, Ni Y, Olsen HS, Rosen CA, Rubin SM,
 CC Shi Y, Siefert TR, Wei Y, Young P, Yu G, Zeng Z,
 CC WPI 93-034967-96
 CC N-PSDB: V84600.
 CC
 CC PI New isolated human genes and the secreted polypeptides they encode -
 CC useful for diagnosis and treatment of e.g. cancers, neurological
 CC disorders, immune diseases, inflammation or blood disorders
 CC Disclosures: Page 147, 722FF, English
 CC The invention relates to nucleic acid sequences (V8441 to V84633)
 CC encoding human secreted proteins (W82034 to W82676). The secreted protein
 CC gene sequences are deposited with the ATCC under deposit numbers ATCC
 CC 97070, 97071, 97072, 97073, 97074, 97075, 97076, 97077, 97078, 97079,
 CC 97080, 97081, 97082, 97083, 97084, 97085, 97086, 97087, 97088,
 CC 97089, 97090, 97091, 97092, 97093, 97094, 97095, 97096, 97097, 97098,
 CC 97099, 97100, 97101, 97102, 97103, 97104, 97105, 97106, 97107, 97108,
 CC 97109, 97110, 97111, 97112, 97113, 97114, 97115, 97116, 97117, 97118,
 CC 97119, 97120, 97121, 97122, 97123, 97124, 97125, 97126, 97127, 97128,
 CC 97129, 97130, 97131, 97132, 97133, 97134, 97135, 97136, 97137, 97138,
 CC 97139, 97140, 97141, 97142, 97143, 97144, 97145, 97146, 97147, 97148,
 CC 97149, 97150, 97151, 97152, 97153, 97154, 97155, 97156, 97157, 97158,
 CC 97159, 97160, 97161, 97162, 97163, 97164, 97165, 97166, 97167, 97168,
 CC 97169, 97170, 97171, 97172, 97173, 97174, 97175, 97176, 97177, 97178,
 CC 97179, 97180, 97181, 97182, 97183, 97184, 97185, 97186, 97187, 97188,
 CC 97189, 97190, 97191, 97192, 97193, 97194, 97195, 97196, 97197, 97198,
 CC 97199, 97200, 97201, 97202, 97203, 97204, 97205, 97206, 97207, 97208,
 CC 97209, 97210, 97211, 97212, 97213, 97214, 97215, 97216, 97217, 97218,
 CC 97219, 97220, 97221, 97222, 97223, 97224, 97225, 97226, 97227, 97228,
 CC 97229, 97230, 97231, 97232, 97233, 97234, 97235, 97236, 97237, 97238,
 CC 97239, 97240, 97241, 97242, 97243, 97244, 97245, 97246, 97247, 97248,
 CC 97249, 97250, 97251, 97252, 97253, 97254, 97255, 97256, 97257, 97258,
 CC 97259, 97260, 97261, 97262, 97263, 97264, 97265, 97266, 97267, 97268,
 CC 97269, 97270, 97271, 97272, 97273, 97274, 97275, 97276, 97277, 97278,
 CC 97279, 97280, 97281, 97282, 97283, 97284, 97285, 97286, 97287, 97288,
 CC 97289, 97290, 97291, 97292, 97293, 97294, 97295, 97296, 97297, 97298,
 CC 97299, 97300, 97301, 97302, 97303, 97304, 97305, 97306, 97307, 97308,
 CC 97309, 97310, 97311, 97312, 97313, 97314, 97315, 97316, 97317, 97318,
 CC 97319, 97320, 97321, 97322, 97323, 97324, 97325, 97326, 97327, 97328,
 CC 97329, 97330, 97331, 97332, 97333, 97334, 97335, 97336, 97337, 97338,
 CC 97339, 97340, 97341, 97342, 97343, 97344, 97345, 97346, 97347, 97348,
 CC 97349, 97350, 97351, 97352, 97353, 97354, 97355, 97356, 97357, 97358,
 CC 97359, 97360, 97361, 97362, 97363, 97364, 97365, 97366, 97367, 97368,
 CC 97369, 97370, 97371, 97372, 97373, 97374, 97375, 97376, 97377, 97378,
 CC 97379, 97380, 97381, 97382, 97383, 97384, 97385, 97386, 97387, 97388,
 CC 97389, 97390, 97391, 97392, 97393, 97394, 97395, 97396, 97397, 97398,
 CC 97399, 97400, 97401, 97402, 97403, 97404, 97405, 97406, 97407, 97408,
 CC 97409, 97410, 97411, 97412, 97413, 97414, 97415, 97416, 97417, 97418,
 CC 97419, 97420, 97421, 97422, 97423, 97424, 97425, 97426, 97427, 97428,
 CC 97429, 97430, 97431, 97432, 97433, 97434, 97435, 97436, 97437, 97438,
 CC 97439, 97440, 97441, 97442, 97443, 97444, 97445, 97446, 97447, 97448,
 CC 97449, 97450, 97451, 97452, 97453, 97454, 97455, 97456, 97457, 97458,
 CC 97459, 97460, 97461, 97462, 97463, 97464, 97465, 97466, 97467, 97468,
 CC 97469, 97470, 97471, 97472, 97473, 97474, 97475, 97476, 97477, 97478,
 CC 97479, 97480, 97481, 97482, 97483, 97484, 97485, 97486, 97487, 97488,
 CC 97489, 97490, 97491, 97492, 97493, 97494, 97495, 97496, 97497, 97498,
 CC 97499, 97500, 97501, 97502, 97503, 97504, 97505, 97506, 97507, 97508,
 CC 97509, 97510, 97511, 97512, 97513, 97514, 97515, 97516, 97517, 97518,
 CC 97519, 97520, 97521, 97522, 97523, 97524, 97525, 97526, 97527, 97528,
 CC 97529, 97530, 97531, 97532, 97533, 97534, 97535, 97536, 97537, 97538,
 CC 97539, 97540, 97541, 97542, 97543, 97544, 97545, 97546, 97547, 97548,
 CC 97549, 97550, 97551, 97552, 97553, 97554, 97555, 97556, 97557, 97558,
 CC 97559, 97560, 97561, 97562, 97563, 97564, 97565, 97566, 97567, 97568,
 CC 97569, 97570, 97571, 97572, 97573, 97574, 97575, 97576, 97577, 97578,
 CC 97579, 97580, 97581, 97582, 97583, 97584, 97585, 97586, 97587, 97588,
 CC 97589, 97590, 97591, 97592, 97593, 97594, 97595, 97596, 97597, 97598,
 CC 97599, 97600, 97601, 97602, 97603, 97604, 97605, 97606, 97607, 97608,
 CC 97609, 97610, 97611, 97612, 97613, 97614, 97615, 97616, 97617, 97618,
 CC 97619, 97620, 97621, 97622, 97623, 97624, 97625, 97626, 97627, 97628,
 CC 97629, 97630, 97631, 97632, 97633, 97634, 97635, 97636, 97637, 97638,
 CC 97639, 97640, 97641, 97642, 97643, 97644, 97645, 97646, 97647, 97648,
 CC 97649, 97650, 97651, 97652, 97653, 97654, 97655, 97656, 97657, 97658,
 CC 97659, 97660, 97661, 97662, 97663, 97664, 97665, 97666, 97667, 97668,
 CC 97669, 97670, 97671, 97672, 97673, 97674, 97675, 97676, 97677, 97678,
 CC 97679, 97680, 97681, 97682, 97683, 97684, 97685, 97686, 97687, 97688,
 CC 97689, 97690, 97691, 97692, 97693, 97694, 97695, 97696, 97697, 97698,
 CC 97699, 97700, 97701, 97702, 97703, 97704, 97705, 97706, 97707, 97708,
 CC 97709, 97710, 97711, 97712, 97713, 97714, 97715, 97716, 97717, 97718,
 CC 97719, 97720, 97721, 97722, 97723, 97724, 97725, 97726, 97727, 97728,
 CC 97729, 97730, 97731, 97732, 97733, 97734, 97735, 97736, 97737, 97738,
 CC 97739, 97740, 97741, 97742, 97743, 97744, 97745, 97746, 97747, 97748,
 CC 97749, 97750, 97751, 97752, 97753, 97754, 97755, 97756, 97757, 97758,
 CC 97759, 97760, 97761, 97762, 97763, 97764, 97765, 97766, 97767, 97768,
 CC 97769, 97770, 97771, 97772, 97773, 97774, 97775, 97776, 97777, 97778,
 CC 97779, 97780, 97781, 97782, 97783, 97784, 97785, 97786, 97787, 97788,
 CC 97789, 97790, 97791, 97792, 97793, 97794, 97795, 97796, 97797, 97798,
 CC 97799, 97800, 97801, 97802, 97803, 97804, 97805, 97806, 97807, 97808,
 CC 97809, 97810, 97811, 97812, 97813, 97814, 97815, 97816, 97817, 97818,
 CC 97819, 97820, 97821, 97822, 97823, 97824, 97825, 97826, 97827, 97828,
 CC 97829, 97830, 97831, 97832, 97833, 97834, 97835, 97836, 97837, 97838,
 CC 97839, 97840, 97841, 97842, 97843, 97844, 97845, 97846, 97847, 97848,
 CC 97849, 97850, 97851, 97852, 97853, 97854, 97855, 97856, 97857, 97858,
 CC 97859, 97860, 97861, 97862, 97863, 97864, 97865, 97866, 97867, 97868,
 CC 97869, 97870, 97871, 97872, 97873, 97874, 97875, 97876, 97877, 97878,
 CC 97879, 97880, 97881, 97882, 97883, 97884, 97885, 97886, 97887, 97888,
 CC 97889, 97890, 97891, 97892, 97893, 97894, 97895, 97896, 97897, 97898,
 CC 97899, 97900, 97901, 97902, 97903, 97904, 97905, 97906, 97907, 97908,
 CC 97909, 97910, 97911, 97912, 97913, 97914, 97915, 97916, 97917, 97918,
 CC 97919, 97920, 97921, 97922, 97923, 97924, 97925, 97926, 97927, 97928,
 CC 97929, 97930, 97931, 97932, 97933, 97934, 97935, 97936, 97937, 97938,
 CC 97939, 97940, 97941, 97942, 97943, 97944, 97945, 97946, 97947, 97948,
 CC 97949, 97950, 97951, 97952, 97953, 97954, 97955, 97956, 97957, 97958,
 CC 97959, 97960, 97961, 97962, 97963, 97964, 97965, 97966, 97967, 97968,
 CC 97969, 97970, 97971, 97972, 97973, 97974, 97975, 97976, 97977, 97978,
 CC 97979, 97980, 97981, 97982, 97983, 97984, 97985, 97986, 97987, 97988,
 CC 97989, 97990, 97991, 97992, 97993, 97994, 97995, 97996, 97997, 97998,
 CC 97999, 98000, 98001, 98002, 98003, 98004, 98005, 98006, 98007, 98008,
 CC 98009, 98010, 98011, 98012, 98013, 98014, 98015, 98016, 98017, 98018,
 CC 98019, 98020, 98021, 98022, 98023, 98024, 98025, 98026, 98027, 98028,
 CC 98029, 98030, 98031, 98032, 98033, 98034, 98035, 98036, 98037, 98038,
 CC 98039, 98040, 98041, 98042, 98043, 98044, 98045, 98046, 98047, 98048,
 CC 98049, 98050, 98051, 98052, 98053, 98054, 98055, 98056, 98057, 98058,
 CC 98059, 98060, 98061, 98062, 98063, 98064, 98065, 98066, 98067, 98068,
 CC 98069, 98070, 98071, 98072, 98073, 98074, 98075, 98076, 98077, 98078,
 CC 98079, 98080, 98081, 98082, 98083, 98084, 98085, 98086, 98087, 98088,
 CC 98089, 98090, 98091, 98092, 98093, 98094, 98095, 98096, 98097, 98098,
 CC 98099, 98100, 98101, 98102, 98103, 98104, 98105, 98106, 98107, 98108,
 CC 98109, 98110, 98111, 98112, 98113, 98114, 98115, 98116, 98117, 98118,
 CC 98119, 98120, 98121, 98122, 98123, 98124, 98125, 98126, 98127, 98128,
 CC 98129, 98130, 98131, 98132, 98133, 98134, 98135, 98136, 98137, 98138,
 CC 98139, 98140, 98141, 98142, 98143, 98144, 98145, 98146, 98147, 98148,
 CC 98149, 98150, 98151, 98152, 98153, 98154, 98155, 98156, 98157, 98158,
 CC 98159, 98160, 98161, 98162, 98163, 98164, 98165, 98166, 98167, 98168,
 CC 98169, 98170, 98171, 98172, 98173, 98174, 98175, 98176, 98177, 98178,
 CC 98179, 98180, 98181, 98182, 98183, 98184, 98185, 98186, 98187, 98188,
 CC 98189, 98190, 98191, 98192, 98193, 98194, 98195, 98196, 98197, 98198,
 CC 98199, 98200, 98201, 98202, 98203, 98204, 98205, 98206, 98207, 98208,
 CC 98209, 98210, 98211, 98212, 98213, 98214, 98215, 98216, 98217, 98218,
 CC 98219, 98220, 98221, 98222, 98223, 98224, 98225, 98226, 98227, 98228,
 CC 98229, 98230, 98231, 98232, 98233, 98234, 98235, 98236, 98237, 98238,
 CC 98239, 98240, 98241, 98242, 98243, 98244, 98245, 98246, 98247, 98248,
 CC 98249, 98250, 98251, 98252, 98253, 98254, 98255, 98256, 98257, 98258,
 CC 98259, 98260, 98261, 98262, 98263, 98264, 98265, 98266, 98267, 98268,
 CC 98269, 98270, 98271, 98272, 98273, 98274, 98275, 98276, 98277, 98278,
 CC 98279, 98280, 98281, 98282, 98283, 98284, 98285, 98286, 98287, 98288,
 CC 98289, 98290, 98291, 98292, 98293, 98294, 98295, 98296, 98297, 98298,
 CC 98299, 98300, 98301, 98302, 98303, 98304, 98305, 98306, 98307, 98308,
 CC 98309, 98310, 98311, 98312, 98313, 98314, 98315, 98316, 98317, 98318,
 CC 98319, 98320, 98321, 98322, 98323, 98324, 98325, 98326, 98327, 98328,
 CC 98329, 98330, 98331, 98332, 98333, 98334, 98335, 98336, 98337, 98338,
 CC 98339, 98340, 98341, 98342, 98343, 98344, 98345, 98346, 98347, 98348,
 CC 98349, 98350, 98351, 98352, 98353, 98354, 98355, 98356, 98357, 98358,
 CC 98359, 98360, 98361, 98362, 98363, 98364, 98365, 98366, 98367, 98368,
 CC 98369, 98370, 98371, 98372, 98373, 98374, 98375, 98376, 98377, 98378,
 CC 98379, 98380, 98381, 98382, 98383, 98384, 98385, 98386, 98387, 98388,
 CC 98389, 98390, 98391, 98392, 98393, 98394, 98395, 98396, 98397, 98398,
 CC 98399, 98400, 98401, 98402, 98403, 98404, 98405, 98406, 98407, 98408,
 CC 98409, 98410, 98411, 98412, 98413, 98414, 98415, 98416, 98417, 98418,
 CC 98419, 98420, 98421, 98422, 98423, 98424, 98425, 98426, 98427, 98428,
 CC 98429, 98430, 98431, 98432, 98433, 98434, 98435, 98436, 98437, 98438,
 CC 98439, 98440, 98441, 98442, 98443, 98444, 98445, 98446, 98447, 98448,
 CC 98449, 98450, 98451, 98452, 98453, 98454, 98455, 98456, 98457, 98458,
 CC 98459, 98460, 98461, 98462, 98463, 98464, 98465, 98466, 98467, 98468,
 CC 98469,



CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 371 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC HYPOTHETICAL: NO
 CC ANTI-SENSE: NO
 CC FRAGMENT TYPE: internal
 CC ORIGINAL SOURCE:
 SO SEQUENCE 371 AA: 42741 MW: 690367 CN:

Query Match 94 98 Score 37 DB 2 Length 371
 Best Local Similarity 90.0% Pos 3 Nc 2172-22
 Matches 4: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Lb 103 VVRPP 107
 : : : :
 Uy 1 VVRPP 5

RESULT 3
 ID US-07-912-190A-2 STANDARD: PRI: 1157 AA.
 XX
 AC xxxxxx
 DT

DE Sequence 2 Application "Scorpion-190A"
 CC Sequence 2, Application US/07912190A
 CC Patent No. 5368993
 CC GENERAL INFORMATION:
 CC APPLICANT: Ponceirada, Luis R
 CC APPLICANT: Ponceirada, Jewel W
 CC APPLICANT: Sick, August J
 CC TITLE OF INVENTION: Baculovirus for Controlling Scarab Pests
 CC TITLE OF INVENTION: thuringiensis isolate and a N. 53689924 Gene Encoding a
 CC TITLE OF INVENTION: Coleopteran-Active Toxin
 CC NUMBER OF SEQUENCES: 2
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Roman Salwanchuk
 CC STREET: 2421 N.W. 41st Street, Suite A-1
 CC CITY: Gainesville
 CC STATE: FL
 CC COUNTRY: USA
 CC ZIP: 32606
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: 07/07812190A
 CC FILING DATE: 19920102
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Salwanchuk, Roman
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 904 375-8100
 CC TELEFAX: 904 372-5900
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1157 amino acids
 CC TYPE: AMINO ACID
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC HYPOTHETICAL: YES
 CC ANTI-SENSE: NO
 CC ORIGINAL SOURCE:
 CC ORGANISM: Bacillus thuringiensis
 CC STRAIN: kumamotoensis

CC INDIVIDUAL ISOLATE: PS50C
 CC IMMEDIATE SOURCE:
 CC LIBRARY: Lambdagem (TM) - 11 LIBRARY OF IUTS
 CC LIBRARY: FONCERRADA
 CC CLONE: 50C
 SO SEQUENCE 1157 AA: 121003 MW: 5374017 CN:

Query Match 94 98 Score 37 DB 1 Length 1157
 Best Local Similarity 90.0% Pos 1 Nc 2172-22
 Matches 4: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 339 VVRPP 343
 : : : :
 Uy 1 VVRPP 5

RESULT 4
 ID US-08-345-468-2 STANDARD: PRI: 1157 AA.
 XX
 AC xxxxxx
 DT

DE Sequence 2 Application "Scorpion-190A"
 CC Sequence 2, Application US/08345468
 CC Patent No. 5554534
 CC GENERAL INFORMATION:
 CC APPLICANT: Nichols, Terry Ellis
 CC APPLICANT: Ponceirada, Luis
 CC APPLICANT: Narva, Kenneth E.
 CC TITLE OF INVENTION: Process for Controlling Scarab Pests
 CC TITLE OF INVENTION: with Bacillus thuringiensis is dates
 CC NUMBER OF SEQUENCES: 6
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: David R. Salwanchuk
 CC STREET: 2421 N.W. 41st Street, Suite A-1
 CC CITY: Gainesville
 CC STATE: FL
 CC COUNTRY: USA
 CC ZIP: 32606
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: 08/08345468
 CC FILING DATE: 19920102
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/08345468
 CC FILING DATE: 01 FEB 1993
 CC APPLICATION NUMBER: 07/929,430
 CC FILING DATE: 30 JAN 1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/909,316
 CC FILING DATE: 16 DEC 1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Salwanchuk, David R.
 CC REGISTRATION NUMBER: 31794
 CC REFERENCE CITE NUMBER: 07/929,430
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 904-375-8100
 CC TELEFAX: 904-372-5900
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1157 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC HYPOTHETICAL: YES

Query Match 94.4% Score 0.000 Length 1227
 Best local similarity 80.0% Pred. No. 2.17e+02
 Matches 4 Conservative 0 Mismatches 0 Indels 0 Gaps 0

DB 325 VWRPP 329
 1.1.1
 0y 1 VWRPP 5

Search completed. Fri Feb 25 13:19:44 2000
 Job time : 45 secs.

[WATERMAN] (M)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1994-1999 University of Pittsburgh, PA
Distributor rights by Oxford Molecular Ltd

Mistchelp protein = protein database search, using Smith-Waterman algorithm
File 103 27 13.15.24 2155, Masfar time 2.33 Seconds
68.411 Million cell updates/sec

Tabular output not generated.

File 103 27 13.15.24 2155, Masfar time 2.33 Seconds
68.411 Million cell updates/sec

File 103 27 13.15.24 2155, Masfar time 2.33 Seconds
68.411 Million cell updates/sec

Searches: 122810 seqs, 4006613 residues

Post-processor: Minimum Match 90%

File 103 27 13.15.24 2155, Masfar time 2.33 Seconds
68.411 Million cell updates/sec

File 103 27 13.15.24 2155, Masfar time 2.33 Seconds
68.411 Million cell updates/sec

File 103 27 13.15.24 2155, Masfar time 2.33 Seconds
68.411 Million cell updates/sec

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Exp. #	Length	HR	ID	Description	Prod. No.
1	46	100.00	172	2	S28082	Vit protein - simian	1.73e-01
2	46	100.00	183	2	S28220	Cruciferin (clon)	1.73e-01
3	46	100.00	196	2	S26218	Cruciferin (clon)	1.73e-01
4	46	100.00	211	2	R25162	3SK prolactin-rich prot	1.73e-01
5	46	100.00	212	2	S07652	Cruciferin (clon)	1.73e-01
6	46	100.00	234	2	S28227	Cruciferin (clon)	1.73e-01
7	46	100.00	235	2	S28260	Prolactin-rich cell wall	1.73e-01
8	46	100.00	242	2	A47148	Hypothetical 32K prot	1.73e-01
9	46	100.00	425	2	A47148	Beta-lactamase (EC 3)	1.73e-01
10	46	100.00	481	2	S52985	Cell wall protein - a	1.73e-01
11	46	100.00	481	2	S64465	Leucine precursor (cl)	1.73e-01
12	46	100.00	474	2	E71419	Probable indole-3-ace	1.73e-01
13	46	100.00	479	2	S28223	Cruciferin precursor	1.73e-01
14	46	100.00	481	2	T01850	UDP-glucose glucosyl	1.73e-01
15	46	100.00	509	2	A45540	Cruciferin 1 precursor	1.73e-01
16	46	100.00	534	2	S41735	Cholesterol esterase	1.73e-01
17	46	100.00	538	2	S05684	Lipase 1 precursor	1.73e-01
18	46	100.00	549	2	JN0553	Triacylglycerol lipas	1.73e-01
19	46	100.00	549	2	S24448	Triacylglycerol lipas	1.73e-01
20	46	100.00	549	2	JN0551	Triacylglycerol lipas	1.73e-01
21	46	100.00	706	2	S62933	Hypothetical protein	1.73e-01
22	46	100.00	773	2	A47666	Structural polyprotein	1.73e-01
23	46	100.00	1224	2	S41199	Probable adenylate/44	1.73e-01

24	39	100.00	186	1	GNWVDF	Cruciferin (clon)	1.73e-01
25	37	94.9	99	2	S23904	Cell wall protein - a	1.73e-01
26	37	94.9	183	2	S20844	Modulation protein in	1.73e-01
27	37	94.9	190	2	S07100	Modulation protein in	1.73e-01
28	37	94.9	233	1	H75666	Probable serine prote	1.73e-01
29	37	94.9	233	1	S28223	Cruciferin (clon)	1.73e-01
30	37	94.9	268	2	T01425	Conserved hypothetical	1.73e-01
31	37	94.9	277	2	S52985	Cell wall protein - a	1.73e-01
32	37	94.9	341	2	S07652	Cruciferin (clon)	1.73e-01
33	37	94.9	371	2	S41199	Probable adenylate/44	1.73e-01
34	37	94.9	371	2	S51320	Mitochondrial protein	1.73e-01
35	37	94.9	484	2	R66233	N-ethylmaleimide-acti	1.73e-01
36	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
37	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
38	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
39	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
40	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
41	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
42	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
43	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
44	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
45	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
46	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
47	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
48	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
49	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
50	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
51	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
52	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01
53	37	94.9	492	2	T01445	Hypothetical protein	1.73e-01

Note: Post-processor removed 947 summaries from list due to search parameters of Seq.

ALIGNMENTS

RESULT	1
ENTRY	S28082
TITLE	Vit protein - simian immunoglobulin virus
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28082
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541
TITLE	Sequence of a novel simian immunoglobulin virus from a wild-caught African mandrill
CROSS-REFERENCES	M010:90015168
ABSTRACT	228584
STATUS	nucleic acid sequence not shown; translation not shown
MOLECULE	1.73e-01
PROTEINS	1.73e-01
NOTES	the nucleotide sequence was submitted to the EMBL data library, July 1989
GENETICS	vit
NAME	1.73e-01
SUMMARY	1.73e-01
Query Match	100.00%
Best Local Similarity	100.00%
Matches	5: Conservative 5: Mismatches 5: Gaps
DB	95 VVRPP 99
CV	1 VVRPP 5
RESULT	2
ENTRY	S28220
TITLE	Cruciferin precursor
ORGANISM	Formal name simian immunoglobulin virus, SIV
DATE	17-Apr-1993
ADDITIONS	17-Mar-1994
REFERENCE	S28220
AUTHORS	Isajima, H.; Hasegawa, A.; Maki, N.; Fukusawa, M.; Matsuoka, T.; Speidel, S.; Cooper, K.W.; Matsuda, E.N.; ...
JOURNAL	Nature (1989) 41:539-541


```

##cross-references EMBL:M36761
##experimental_source strain Ref S2392, clone pM551
##note the authors designated the arg codon for residue 1 as
start codon: the GTG codon for residue 1 was
translated as Val

#accession S27006
#molecule_type protein
#residues 47-57189 96 #label 152
#experimental_source overexpressed cloned gene from strain F02-S-962
REFERENCE S01468
#authors de Weester, P., Joris, B., Lenzi, M.V., Delottay, P.,
Erpicum, T., Desart, Y., Klein, D., Ohysen, M., Proye,
J.M., van Beumen, J.
#journal Biochem. J. (1981) 241:427-432
#title The active sites of the beta-lactamase of Streptomyces
cacaoli and Streptomyces albus 3.
#cross-references MIM:581411
#accession S11762
#molecule_type protein
#residues 89-96 #label DEM
#experimental_source cloned gene from plasmid pMX51 overexpressed in
Streptomyces lividans
COMMENT This is a class A beta-lactamase.
CLASSIFICATION #superfamily beta-lactamase I
KEYWORDS antibiotic resistance; extracellular protein; hydrolase
FEATURE
1-45 #domain signal sequence #status predicted #label S15X
47-245 #product beta-lactamase #status experimental #label MAT,
93 #active-site Ser #status experimental
SUMMARY #length 325 #molecular_weight 37557 #checksum 2127
Query Match 100.0% Score 39.18 #length 475
Best Local Similarity 100.0% Pred No. 1.73e-015
Matches 5. Conservation 0. Mismatches 0. Indels 0. Gaps 0.

El 291 VRRFP 239
|||||
QY 1 VRRPP 5

RESULT 10
ENTRY S52985 #type complete
TITLE Cell wall protein - alfalfa
#formal_name Medicago sativa #common_name alfalfa
ORGANISM 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change
04-Sep-1998
ACCESSIONS S52985
REFERENCE S52985
#authors Heath, G.E., Winicov, I.
#journal Plant Mol. Biol. (1985) 27:411-418
#title Post-transcriptional regulation of a salt-inducible alfalfa
gene encoding a putative chitin-binding protein
#cross-references MIM:5619167
#accession S52985
#status preliminary
#molecule_type RNA
#residues 1-381 #label pM
#cross-references EMBL:14017 NID-9556120 F01056413
CLASSIFICATION #superfamily beta-lactamase I
SUMMARY #length 392 #molecular_weight 47557 #checksum 4221
Query Match 100.0% Score 39.18 #length 381
Best Local Similarity 100.0% Pred No. 1.73e+01
Matches 5. Conservation 0. Mismatches 0. Indels 0. Gaps 0.

Db 146 VRRPP 150
|||||
QY 1 VRRPP 5

RESULT 11
ENTRY S68963 #type complete

```

```

#formal_name Magnolia salicifolia
#formal_name Magnolia salicifolia
DATE 14-Mar-1998 #sequence_revision 14-Mar-1997 #text_change
04-Sep-1997
ACCESSIONS S68963; S77067; S54308
REFERENCE S68964
#authors Fischer, H., Haake, V., Horstmann, C., Jensen, U.
#journal Eur. J. Biochem. (1995) 229:645-650
#title Characterization of a cDNA encoding a highly related protein of Magnolia
legumin-encoding cDNAs representing two divergent gene
subfamilies.
#cross-references MIM:5670215
#accession S68965
#molecule_type mRNA
#residues 1-474 #label F151
#cross-references MIM: X82464 MIM:570567 F11 570566
#note the authors translated the cDNA cDNA for residue 462 as
Val
#accession S77067
#molecule_type protein
#residues 462 #label F152
CLASSIFICATION #superfamily glycinin
KEYWORDS seed; storage protein
FEATURE
1-23 #domain signal sequence #status predicted #label S15X
24-299 #product legumin alpha chain #status predicted #label
299-476 #product legumin beta chain #status experimental #label
MAT12
SUMMARY #length 477 #molecular_weight 57000 #checksum 7790
Query Match 100.0% Score 39.18 #length 475
Best Local Similarity 100.0% Pred No. 1.73e-015
Matches 5. Conservation 0. Mismatches 0. Indels 0. Gaps 0.

El 270 VRRFP 274
|||||
QY 1 VRRPP 5

RESULT 12
ENTRY E71419 #type complete
TITLE Probable indole-3-acetate beta-glucosyltransferase -
Arabidopsis thaliana
#formal_name Arabidopsis thaliana #common_name mouse-ear
cress
ORGANISM Columbia
#variety Columbia
DATE 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change
05-Feb-1998
ACCESSIONS E71419
REFERENCE A71400
#authors Pao, M., Pao, F., Pao, F., Pao, F., Pao, F., Pao, F.,
Dean, C., Bergkamp, R., Dirks, W., Van Staveren, M.,
Stiekema, W., Frost, L., Ridley, P., Hudson, S.A., Patel,
P., Murphy, P., Piffanelli, P., Wedler, H., Wedler, E.,
Wambatta, R., Wittenberger, R., Pohl, T., Terry, N.,
Gleason, J., Villarroel, R., De Clerck, R., Van Montagu, M.,
Lechardy, A., Aubert, S., Gy, I., Kreis, M., Lao, N.,
Kavanagh, T., Hempel, S., Kottler, P., Entian, K.D., Rieger,
M., Schaffner, W., Fink, E., Meller, A., S. Silver, M.,
Tardieu, P., Maffei, A., Pao, F., Pao, F., Pao, F., Pao, F.,
A. Voulalas, E., Milina, D., Hatzopoulos, P.,
Piravandi, E., Obermaier, B., Hilbert, H., Duesterhoff, A.,
Mooros, T., Jones, T.D., Fieva, T., Palmer, K., Benes, V.,
Kutman, S., Acosta, W., Cacho, P., Berger, C., Delsen,
M., Voet, M., Volckaert, G., Mewes, H.W., Klosterman, S.,
Schueller, C., Chalwatzis, N.
#journal Nature (1998) 391:485-488
#title Analysis of 1.4 Mb of contiguous sequence from chromosome 4
of Arabidopsis thaliana.
#cross-references MIM:68121113
#accession E71419
#status preliminary; nucleic acid sequence not shown

```


Query Match 100.0%; Score 39; DB 2; Length 509;
 Best Local Similarity 100.0%; Pred No 1 71e-01;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 292 VRRP 296
 |||||
 Qy 1 VRRP 5

Search completed. Fri Feb 25 14:06:56 2000
 Job time : 49 secs.


```

DE PROLINE RICH 35 KD EXTENSIN-RELATED PROTEIN PRECURSOR (FRAGMENT)
US LADUS CARCINIA (CARROT)
OC EUCARYOTA; VIRIDIPHYTES; STREPTOPHYTES; EMEROPHYTES; TRACHEOPHYTES;
OC EUBYLIDITES; SERRATOPHYTES; MAGNOLIOPHYTES; EUPHYTIUMS;
RN ARABIDOPSIS; ARABIDOPSIS; ARABIDOPSIS; ARABIDOPSIS
RP SEQUENCE FROM N.A.
RA CHEN J., VARNER S.E.
RI "Isolation and characterization of cDNA clones for carrot extensin
RI 3.3 kDa protein, a member of the extensin family."
RI PROC NATH. ACAD. SCI. U.S.A. 82:4439-4443(1985)
CI "SIMILARITY: BELONGS TO THE EXTENSIN FAMILY OF EXTENSINS"
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch
CC or send an email to license@isb-sib.ch.)
CC -----
DE EMBL: M1222; GI:67552;
DE PIR: P2162; P2162;
DE CELL WALL; REPEAT; SIGNAL
DE NON-TER
DE SIGNAL
DE CHAIN 15 211
DE POTENTIAL
DE PROLINE RICH 35 KD EXTENSIN-RELATED
DE PROTEIN
DE VARIANT 29 29
DE VARIANT 32 32
DE VARIANT 33 33
DE SEQUENCE 211 AA; 21601 MW; 5949265F QPC30;
Query Match 100.00; Score 29; DB 1; Length 211.
Best Local Similarity 100.00; Prid No 5 46e-00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 184 VWRPP 188
|||||
QY 1 VWRPP 5
RESULT 3
ID COAT-22EV STANDARD; PRT: 212 AA.
AC P14849;
DI 01-APR-1990 (REL 14, CREATED)
DI 01-APR-1990 (REL 14, LAST SEQUENCE UPDATE)
DI 01-MAR-1992 (REL 22, LAST ANNOTATION UPDATE)
DE COAT PROTEIN (CAPSID PROTEIN)
OS PEA EARLY BROWNING VIRUS
OC VIRUS; SSRNA POSITIVE-STRAND; VIRUSES; NO DNA STAGE; TORREAVIRUS
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SP5;
RA MEDLINE: 91370950;
RA PETERSEN S.G., IERHARDT J., ROEPHART P.;
RI "Analysis of RNA of pea early browning virus strain SP5."
RI PLANT MOL. BIOL. 13:735-737(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SP5;
RA MEDLINE: 90356386;
RA GOUNON M.G., LOMONOSOFF G.P., DAVIES J.W., WOOD K.R.;
RI "The complete nucleotide sequence of PEBV RNA2 reveals the presence
RI of a novel open reading frame and provides insights into the
RI structure of toroviral subgenomic promoters."
RI NUCLEIC ACIDS RES. 18:4507-4512(1990).
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch
CC or send an email to license@isb-sib.ch.)
CC -----

```

```

CC or send an email to license@isb-sib.ch.)
CC -----
DE EMBL: X15883; GI:20811;
DE PIR: X61928; X61928;
DE PIR: S07552; S07552;
DE PIR: S14698; S14698;
DE COAT PROTEIN
DE SEQUENCE 212 AA; 22244 MW; 11003995 QPC32;
Query Match 100.00; Score 39; DB 1; Length 212.
Best Local Similarity 100.00; Prid No 5 46e-00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 161 VWRPP 165
|||||
QY 1 VWRPP 5
RESULT 5
ID REAL-STREI STANDARD; PRT: 325 AA.
AC Q03680;
DI 01-FEB-1994 (REL 28, CREATED)
DI 01-FEB-1994 (REL 28, LAST SEQUENCE UPDATE)
DI 15-FEB-1998 (REL 3, LAST ANNOTATION UPDATE)
DE REAL-CATINASE 1 PRECURSOR (CD 3.1.2.2) (PENCILLIUM)
GN BLAL;
OS STREPTOMYCES CACAOI;
OC BACTERIA; FRIMUTICES; ACTINOBACTERIA; ACTINOMYCETIDAE;
OC ACTINOMYCETALES; STREPTOMYCETACEAE; STREPTOMYCETIDAE; STREPTOMYCETIDAE;
Query Match 100.00; Score 39; DB 1; Length 292.
Best Local Similarity 100.00; Prid No 5 46e-00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 161 VWRPP 165
|||||
QY 1 VWRPP 5
RESULT 4
ID YHT-ECOLI STANDARD; PRT: 292 AA.
AC P32141;
DI 01-OCT-1993 (REL 27, CREATED)
DI 01-OCT-1993 (REL 27, LAST SEQUENCE UPDATE)
DI 01-NOV-1997 (REL 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 22 KD PROTEIN IN GLNA-SEN INTERGENIC REGION (F032)
GN YHT;
OS ESCHERICHIA COLI;
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RC MEDLINE: 93347969;
RA FLORETTI S., ILLI RUPAND V.D., DANFELS D., RIATNER P.;
RI "Analysis of the Escherichia coli genome. II. DNA sequence of the
RI region from 87.2 to 89.2 minutes."
RI NUCLEIC ACIDS RES. 21:3391-3398(1993).
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch
CC or send an email to license@isb-sib.ch.)
CC -----
DE EMBL: I12201; GI:30426;
DE EMBL: AB004464; GI:740314;
DE PIR: S40925; S40925;
DE ECKHART F.G.I.1946; YHT
DE HYPOTHETICAL PROTEIN.
DE SEQUENCE 292 AA; 31983 MW; F0D184F1 QPC32;
Query Match 100.00; Score 39; DB 1; Length 292.
Best Local Similarity 100.00; Prid No 5 46e-00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 161 VWRPP 165
|||||
QY 1 VWRPP 5
RESULT 5
ID REAL-STREI STANDARD; PRT: 325 AA.
AC Q03680;
DI 01-FEB-1994 (REL 28, CREATED)
DI 01-FEB-1994 (REL 28, LAST SEQUENCE UPDATE)
DI 15-FEB-1998 (REL 3, LAST ANNOTATION UPDATE)
DE REAL-CATINASE 1 PRECURSOR (CD 3.1.2.2) (PENCILLIUM)
GN BLAL;
OS STREPTOMYCES CACAOI;
OC BACTERIA; FRIMUTICES; ACTINOBACTERIA; ACTINOMYCETIDAE;
OC ACTINOMYCETALES; STREPTOMYCETACEAE; STREPTOMYCETIDAE; STREPTOMYCETIDAE;

```

[illegible]

PLANT MOL. BIOL., 20:467-479(1992).

-- FUNCTION: THIS IS A SEED STORAGE PROTEIN.

-- SOURCE: HEMAKER, EMBL SEQUENTIAL IS COMPOSED OF AN ALPHA AND A BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A DISULFIDE BOND.

-- SIMILARITY: BELONGS TO THE L18 SEED STORAGE PROTEINS (L18-FAMILY).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL laboratories at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as it's referred to in the way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (<http://www.ebi.ac.uk/edl/licenses.html>) or send an email to license@ebi.ac.uk.

EMBL: X59408; G21118; -
PIR: S26223; S26224;
PROSITE: PS01351; L18 SEED STORAGE; 1.
PFAM: PF00190; SeedStorage1; 1
SEPT: STORAGE PROTEIN SIGNAL MULTIPRE FAMILY.
SIGNAL: 1 24 HY SIMILARITY:
E1 CHAIN 24 479 CRUCIFERIN PROPEPSEC ALFA BEAN.
E1 CHAIN 240 473 CROCEFERIN PROPEPSEC BETA BEAN.
FT DISULFID 121 206 INTERCHAIN (ALPHA BETA) (POLENTIAL).
E1 DOMAIN 121 141 SENKENSIN-Y-PH.
FT DOMAIN 211 218 POLY-SIN.
SPOTYPE 479 AA 20% IDENTITY WITH PDB 1A90 OR 1A91.

Curry Match: 100.0% Score 36; DR 1; Length 479
Best Local Similarity: 100.0; Field No. 54666
Matches: 1; Complementarity: 1; Mismatches: 1; 11 base

Db 262 VPRFP 266
QY : VPRFP 5

RESULT 7
ID CRU3 PRANA STAN AM: 191; 500 AA.
AC P34525;
DT 01-FEB-1994 (REV. 24, CRITICAL)
UT 01-FEB-1994 (REV. 24, LAST SEQUENCE UPDATE)
DI 01-JUN-1994 (REV. 23, LAST ANN TATION UPDATE)
DE CRUCIFERIN (SEED PROTEIN) (L18 GLIOPOLIN) (L18 STORAGE PROTEIN)
CRUI
OS BRASSICA NAPUS (RAPE)
OC EUPHYASIA VERBESCENTANS (STREET LIME); PHYTA LIMBAE PHYTA LEACHII PHYTA
OC EUPHYATYPETIS SYRIACA PHYTA MAINGII PHYTA RUTICATA PHYTA RUTICATA
OC CAPRARALES: BRASSICACEAE: BRASSICA.
RN 111
RS SEQUENCE FROM N.A.
RA MEDLINE: 93043046.
RT "Characterization of a Brassica napus seed-mediated protein." [Abstract]
RL Submit: estimation of sizes of structural gene families.
RL PLANT MOL. BIOL., 20:469-474(1992).
-- FUNCTION: THIS IS A SEED STORAGE PROTEIN.
-- SOURCE: HEMAKER, EMBL SEQUENTIAL IS COMPOSED OF AN ALPHA AND A BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A DISULFIDE BOND.

-- SIMILARITY: BELONGS TO THE L18 SEED STORAGE PROTEINS (L18-FAMILY).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL laboratories at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as it's referred to in the way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (<http://www.ebi.ac.uk/edl/licenses.html>) or send an email to license@ebi.ac.uk.

EMBL: X56120; G17811; -

```

KW      3D-STRUCTURE.
FT SIGNAL          1      15
FT CHAIN           16     549
FT ACT_SITE        224    549
BY SIMILARITY.
FT ACT_SITE        356    556
BY SIMILARITY.
FT ACT_SITE        464    464
BY SIMILARITY.
FT ACT_SITE        480    480
BY SIMILARITY.
FT ACT_SITE        520    520
BY SIMILARITY.
FT CAPECHUPP       339    339
POTENTIAL.
FT CAPECHUPP       486    486
POTENTIAL.
SEQUENCE 549 AA; 52954 MW; 5540000 CP(30);

Query Match          100%   Score 307.16 kJ Length 549;
Sequence Similarity 100%   E-Val No 5.46e-00;
Matches 57 Conservative 0 Mismatches 0 Indels 0 Gaps 0

Db      117 VWRPP 121
      |||||
QY      1 VWRPP 5

RESULT 9
ID LIPCANCS STANDARD; PRI 549 AA.
AC P00261;
CD 01-FEB-1991 (REL 17, CREATED)
DD 01-SEP-1993 (REL 27, LAST SEQUENCE UPDATE)
EE 15-DEC-1998 (REL 37, LAST ANNOTATION UPDATE)
DE LIPASE 1 PRECURSOR (EC 3.1.1.3).
GN LIPI.
OS CANADIAN REDSEA YEASTS CANDIDA TETRANEXA.
SR EUROPEAN JOURNAL OF CELL BIOLOGY, HEMASCOMYCETES, SACCHAROMYCOTALES,
OC CANDIDACEAE: CANDIDA.
[1]
RP SEQUENCE FROM N.A.
RF STRAIN=ATCC 14830;
RX MEDLINE: 92305068.
RA LONGHI S., FUSEINI F., GRANDOFFI R., ZOTTI M., VANDRI M.,
RA AIREPOGHINA I.;
PT "Candida rugosa lipase: purification of two lipases from the Candida
PT cylindraceae";
RN BYTHIM BIOPHYS ACTA 117: 227-232(1992).
EF [2]
FEF SEQUENCE OF 12,542 FRAM N.A. AND PARTIAL SEQUENCE.
FX STRAIN=ATCC 14830 / MS-S;
FCX MEDLINE: 89364874.
RA KAWAGUCHI Y., KONDA H., TANIGUCHI-MOZUMURA Y., IWASAKI S.;
PT "The cys26c gene is responsible for an aspartic proteolytic yeast Candida
PT cylindraceae";
RL NATURE 341:164-166(1989).
[3]
RFP X-RAY CRYSTALLOGRAPHY (2.06 ANGSTROMS).
MELINE: 92386131.
RA GRONLUND P., LI Y.-N., SCHWAB J.D., BOOTHILLIER F., SMITH P.;
RA HARRISON D., RUBIN E., CYGLER M.;
PT "Analysis of the active site of the lipase from the filamentous fungus
PT Candida rugosa lipase";
RI FEBS LETT 335:199-204(1993).
[4]
RN X-RAY CRYSTALLOGRAPHY (2.00 ANGSTROMS).
MELINE: 94190967.
RA SCHWENK T., BOOTHILLIER F., KOLACHIAS P., REBERT A.N.;
FA STENG P.F., ZIMMER H., CYGLER M.;
PT "Analogs of reaction intermediates identify a unique substrate
PT binding site in Candida rugosa lipase.";
RL BIOCHEMISTRY 33:3434-3450(1994).
[5]
RPP REVIEW
MPDLNP 98451816.
RA BENJAMIN S., PANDEY A.;
PT "Characterization of a novel lipase homologous to Lipo IV and versatility in
PT biotechnology.";
RL YEAST 14:1069-1087(1998).

```


Fri Feb 25 14:06:58 2000

US-09-132-799-2.rsp

Page 8

Search completed: Fri Feb 25 13:17:15 2000
Job time : 47 secs.

A handwriting practice sheet for the letter 'Z'. It features six rows of dashed 'Z' characters on a lined background. The first row is labeled 'Z' and 'z' at the beginning. The second row is labeled 'Z' and 'z' at the beginning. The third row is labeled 'Z' and 'z' at the beginning. The fourth row is labeled 'Z' and 'z' at the beginning. The fifth row is labeled 'Z' and 'z' at the beginning. The sixth row is labeled 'Z' and 'z' at the beginning.

© 1998 by John F. Collins, Biocomputing Research Unit.
Copyright (c) 1998-1999 University of Edinburgh, UK.
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or by any information storage or retrieval system, without prior permission in writing from the copyright owner.

Other three protein-protein database search, using Smith-Waterman algorithm
Score: 67.08 **E-value:** 1E-19 **Seq. Id.** 100% **%Ident.** 100%
Accession G02553 **RefSeq** NP_031734.2(2006) **Maspac time** 4.44 seconds
Database current and nonredundant **Size** 61,508 Million cell updates/sec

[illegible]

Scholarship editor: Robert C. Marsh

SOL RESULTS

[illegible][illegible]

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Protein	N ₁	Seems	814-5	4049	HR	ID	Description	Prod. No.
1	24	24	13471	4	6	029455	UNKNOWN PROTEIN (FRAGMENT)	8.838-00
2	24	24	13471	92	2	007448	CREA.	8.838-00
3	24	24	13471	1	5	042181	CROCIIFERIN TYPE C (FRA	8.838-00
4	24	24	13471	116	5	014685	F21195.8 PROTEIN.	8.838-00
5	24	24	13471	145	4	007404	MAV445.	8.838-00
6	24	24	13471	124	10	002703	CROCIIFERIN (128 STORAG	8.838-00
7	24	24	13471	176	10	042165	CROCIIFERIN (FRAGMENT).	8.838-00
8	24	24	13471	210	14	056247	COAT PROTEIN.	8.838-00
9	24	24	13471	242	14	084598	CHAT PROTEIN.	8.838-00
10	24	24	13471	233	10	041564	CROCIIFERIN (FRAGMENT).	8.838-00
11	24	24	13471	235	10	036496	PROLINE RICH PROTEIN.	8.838-00
12	24	24	13471	249	10	024102	MTN4 (FRAGMENT).	8.838-00
13	24	24	13471	273	10	092371	EVUATIVE MEMBRANE GLYC	8.838-00
14	24	24	13471	420	2	072344	WPKK.	8.838-00
15	24	24	13471	437	2	032655	AMINOMETHYLTRANSFERASE	8.838-00
16	24	24	13471	471	10	093771	GIRINERLIN C20 OXIIA	8.838-00
17	24	24	13471	481	10	040486	CELL WALL PROTEIN.	8.838-00
18	24	24	13471	497	5	022457	11485.7 PROTEIN.	8.838-00
19	24	24	13471	499	14	039246	HYPOTHEICAL 45.5 KD P	8.838-00
20	24	24	13471	421	5	020481	SIMILAR TO POTASSIUM C	8.838-00

Note. Post processing code and a summary of results are available at <http://www.elsevier.com/locate/jmbs>.

APPENDIX

RESULT 1
 029255 PRELIMINARY: 1PT: 6 AA.
 AC 029255;
 EDI 01 NOV-1996 (TREMURER, 01, CREATED)
 DT 01 NOV-1996 (TREMURER, 01, LAST SEQUENCE UPDATE)
 DT 01 MAY-1997 (TREMURER, 03, LAST ANNOTATION UPDATE)
 UNKNOW PROTEIN (TREMURER)
 SUS SCORHA (PIG)
 Q5 EDKARYOTA; METAZA; A; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC ARTHRODACTYLA; SUITA; MAMM; SUINA; SUIDAE; SUI.
 ERN (1)
 SEQUENCE FROM N.A.
 RC TISSUE-SHALL INTEREST:
 RC WINIFEOE A.K., FREDERICK M., DAVIES W.:
 MAMM. GENOME 7:549-557 (1996).
 DR EXBL: F15071: E1967457
 FT NON-TER 1
 FT NON-TER 63
 SEQUENCE 63 AA: 6443 MW: 6882707.0
 Every March 190-09; Score 39; LR 64; Length 63
 Post-Local Similarity 100%; Pred. No. P-44000
 Database 1. Consensus 1.0. Misstatements 0.0. Gaps 0
 DR 51 VVR04 05

```

|||||
QY 1 VVRPP 5

RESULT 2
ID 007848
AC 007848
DT 01-NOV-1997 (PREMREL 04, CREATED)
DI 01-JUL-1997 (TREMREL 04, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (PREMREL 02, LAST ANNOTATION UPDATE)
DE 01-NOV-1998 (PREMREL 02, LAST ANNOTATION UPDATE)
GN CPCR
OS ESCHERICHIA COLI
OC BACTERIA, PROTEOBACTERIA: GAMMA SUBDIVISION: ENTEROBACTERIAEAL
OC ESCHERICHIA
RN (1)
RP SEQUENCE FROM N.A.
PV WFLINE, 0700112
RA HU Y H, HU F, DEAN R, GINZAS M, PETRAFF W, TEIN N J
PT "Genetic analysis of chromosomal islands in Escherichia coli"
RI Chromosome Organization in Escherichia coli
RL GENETICS 143:1521-1532(1996)
DR EMBL: S83896; E33898
SQ SEQUENCE 82 AA, 1004 MW, P29543, G6012.

Query Match 100.00, Score 34, 28.5, Length 126;
Best Local Similarity 100.00; Pred No. 8.84e+00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 VVRPP 39
|||||
QY 1 VVRPP 5

RESULT 3
ID 042181
AC 042181
DI 01-NOV-1998 (PREMREL 01, CREATED)
DI 01-NOV-1998 (PREMREL 01, LAST SEQUENCE UPDATE)
DI 01-AUG-1998 (PREMREL 07, LAST ANNOTATION UPDATE)
DE CROCIFERIN TYPE C (FRAGMENT)
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS)
OC EUPHYLLIPHYTES: SPERMATOPHYTES: MAGNOLIOPHYTES: POLYGYNELES: ROSALES
OC CAPRIFALES: BRASSICACEAE: ARABIDOPSIS
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-DRY SEEDS OF A THAILIAN ECOTYPE COLUMBIA;
RA RAYAL M, GRELLET F, LACIE M, MEYER Y, SOORE R, BELSEN M,
KL SUBMITTED (NOV-1998) TO EMBL/GENBANK/CCP DATA BANKS
DR EMBL: 047512
PT NON-TER
FT NON-TER
SQ SEQUENCE 115 AA, 13019 MW, R10066, G6032.

Query Match 100.00, Score 33, 18.12, Length 115;
Best Local Similarity 100.00; Pred No. 8.84e+00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 84 VVRPP 88
|||||
QY 1 VVRPP 5

RESULT 4
ID 019685
AC 019685
DI 01-NOV-1998 (PREMREL 01, CREATED)
DI 01-NOV-1998 (PREMREL 01, LAST SEQUENCE UPDATE)
DI 01-JAN-1999 (TREMREL 04, LAST ANNOTATION UPDATE)
DE F2105 8 PROTEIN
GN F2105 8
OS CAENORHABDITIS MELANIS.

```

```

OC EUKARYOTA: METAZOA: NEMATODA: SECERNENTEA: PHARODITIDA:
OC PHARODITIDA: PHARODITIDAE: PELICERINAE: CAENORHABDITIS.
RN (1)
RP SEQUENCE FROM N.A.
RA BERKS M J
PT "Sequencing of the 12S storage protein gene from the nematode Caenorhabditis elegans"
RN (2)
RP SEQUENCE FROM N.A.
RA MEDLINE: 94150718
RA WILSON P, AINSWORTH P, ANDERSON K, RAYNES C, BERKS M,
RA ROSEFIELD J, BURTON J, CONNELL M, OUSSEY I, ANDERSON J, GUNSON A,
RA CRAXTON M, DEAR S, DE J, EUBIN R, FAVELLO A, FULTON L,
RA GARDNER A, GREEN P, HAWKINS J, HILLIER N, JEFF M, JERNSTON L,
RA JONES M, KERSHAN J, KIRSTEN J, LAISTER N, PATRILL P,
RA LIGHTING J, LLOYD C, MCMURRAY A, MORTIMORE B, O'CALLAGHAN M,
RA PARSONS J, PERCY C, RIFEN L, ROOPRA A, SAUNDERS D, SHOWNKEEN P,
RA SMALDON N, SMITH A, SONNHAMMER E, STADEN P, SUSTON J,
RA TERRY-MEE J, THOMAS P, VAUGHAN K, WATERSTON P,
RA WATSON A, WELSHAM L, WILKINSON-SPEAR J, WHITEMAN P,
PT "Cloning of the 12S storage protein gene from chromosome III of C. elegans"
RI NATURE 368:32-38(1994)
DR EMBL: Z54271; E1348020
SQ SEQUENCE 105 AA, 11502 MW, F56080, G6020.

Query Match 100.00, Score 33, 28.5, Length 126;
Best Local Similarity 100.00; Pred No. 8.84e+00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 109 VVRPP 113
|||||
QY 1 VVRPP 5

RESULT 5
ID 007404
AC 007404
DI 01-NOV-1997 (PREMREL 04, CREATED)
DI 01-JUL-1997 (PREMREL 04, LAST SEQUENCE UPDATE)
DI 01-NOV-1998 (PREMREL 06, LAST ANNOTATION UPDATE)
DE MAV145
GN MAV145
OS MYCOBACTERIUM AVIUM
OC BACTERIA: FIRMICUTES: ACTINOBACTERIA: ACTINOBACTERIDAE
OC ACTINOMYCETIALES: CORYNEBACTERIINAE: MYCOBACTERIACEAE: MYCOBACTERIUM
RN (1)
RP SEQUENCE FROM N.A.
RC SPAIN-GIP10;
RA MEDLINE: 93195739
RA LAM M, GUSHEV P, DE POST F, SPITALF P, RUTARTI G,
PT "Identification of a 145 kDa variable sequence around the inhA gene
of Mycobacterium avium and similarity analysis of the products of
putative ORFs"
RI MICROBIOLOGY 144:907-914(1998)
DR EMBL: AF002230; G189232
SQ SEQUENCE 145 AA, 15550 MW, E55216, G6022.

Query Match 100.00, Score 33, 28.2, Length 145;
Best Local Similarity 100.00; Pred No. 8.84e+00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 64 VVRPP 68
|||||
QY 1 VVRPP 5

RESULT 6
ID 002503
AC 002503
DI 01-NOV-1998 (PREMREL 01, CREATED)
DI 01-NOV-1998 (PREMREL 01, LAST SEQUENCE UPDATE)
DI 01-NOV-1998 (PREMREL 06, LAST ANNOTATION UPDATE)
DE CROCIFERIN (12S STORAGE PROTEIN; (FRAGMENT))

```


[illegible]

```

DB      51 WVRPP 55
          IIIII
          00000
          1 WVRFP 5
          -----
RESULT   13
-----
ID       CB2271    PRELIMINARY.     FET.    274 AA.
AC       OB2371:
DT       DT 01-NOV-1968 (PEMRELF, OR CREATED)
DT       DT 01-NOV-1968 (PEMRELF, OR LAST SEQUENCE UPDATE)
DT       DT 01-NOV-1968 (PEMRELF, OR LAST ANNOTATION UPDATE)
DE       DE POLYALIVE MEMBRANE GLYCOPROTEIN.
GN       GN T27A16.5.
OS       OS ARABIDOPSIS ITALIANA (MOUSE-EAR CRESS).
OS       OS EMBRYONIA VITISPLANTAE STEPHENOBYTIA EMBRYOBYTIA TRACHEOPYTIA.
OS       OS EPHYLLIOHYLE SPERMATOPHYTIA MAGNOLIOPHYTIA EUCOTYLEDONS FORSIDAE
OS       OS CAPSAFALES; BRASSICACEAE; ARABIDOPSIS.
[1]
RN       RN SEQUENCE FROM N.A.
RP       RP STRAIN=IV COLUMBIA;
RC       RC PENINSLEY D. LIN X KALL S.D. SHUA T.E. FUJII G.Y. MASON J.M.
RPA      RPA SHEN M. FINNINGER M. TRANSFER M.M. SUMERVILLE L.P. VENTER T.C.
RPT      RPT "Arabidopsis thaliana chromosome II RAC T27A16 genomic sequence."
RF       RF SUBMITTED (SEP-1998) TO EMBL/GENBANK/JURY DATA BANKS.
RFL      RFL EMBL AC006496 G3582319;
RFE      RFE GENBANK U000000000;
RSE      RSE SEQUENCE 274 AA. 2799g MW 271367.75 Da.

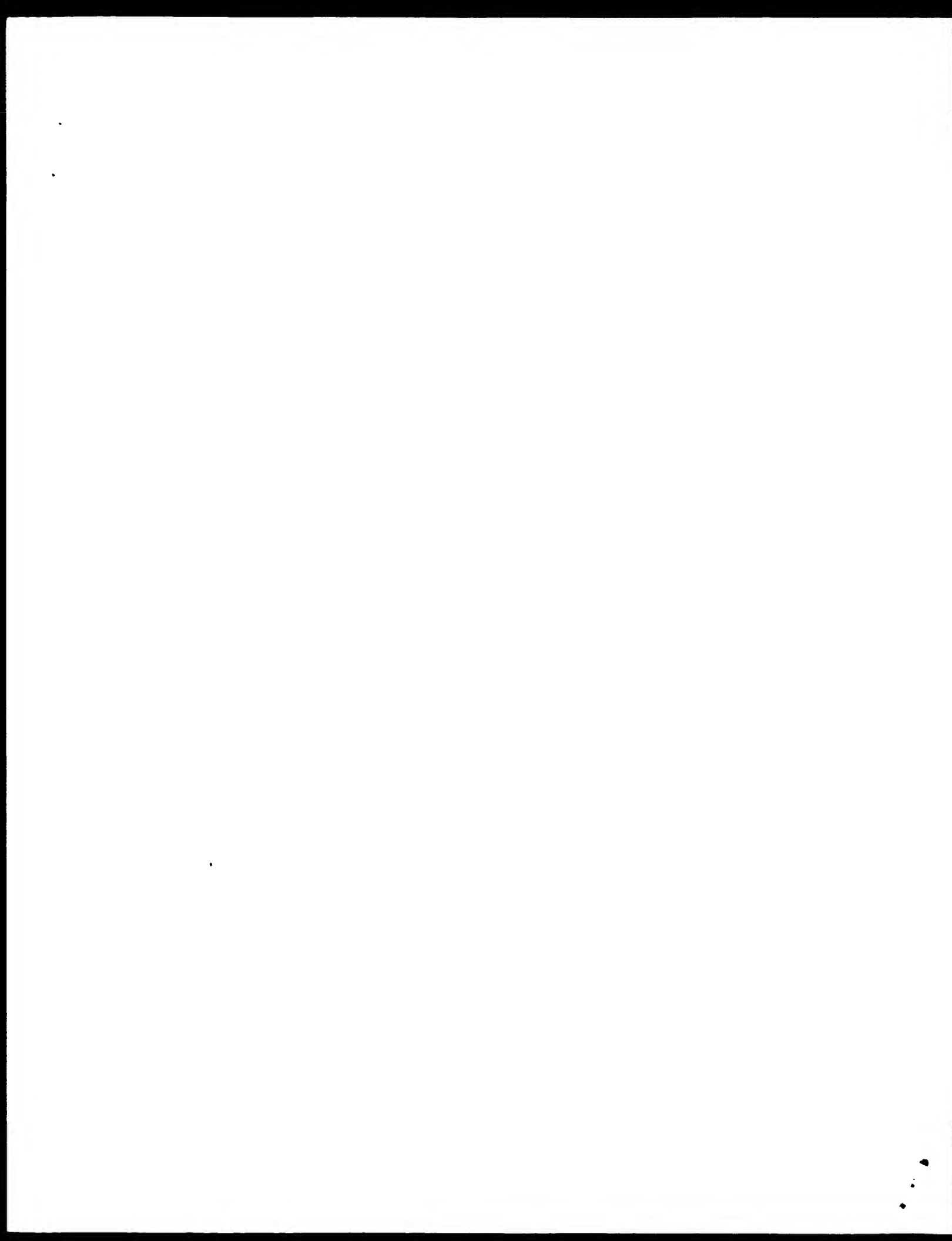
```

```

DE      175  WRPP 179
      11111
QY      1  WRPP 5

      RESULT 14
      ID P72144      PRELIMINARY.      PET. 320 AA.
      AC P72144
      DT 01-FEB-1997 (TEMPREL. 02 CREATED)
      DT 01-FEB-1997 (TEMPREL. 02 LAST SEQUENCE UPDATE)
      CN CNV.1348 (TEMPREL. 05 LAST ANNOTATION UPDATE)
      DE WBRK.
      GN WBRK.
      OS PSEUDOMONAS AERUGINOSA.
      BA BACTERIA: FRIGIDILACTERIA.
      CC PSEUDOMONAS.
      RN [1]

```

Fri Feb 25 14:07:03 2000

US-09-132-799-3.rag

Page 2

Search completed: Fri Feb 25 14:07:03 2000
Job time : 80 secs.




```

NEW FAS
SUMMARY
length 440 #molecular-weight 47657 #checksum 7527
Query Match 95.79 Score 45 PFC length 440:
Best Local Similarity 83.88 PFC No 5,480-00:
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 35 VWRPP 40
|||||
QV 1 VWRPP 6

RESULT 6
ENTRY
TITLE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana #Simulated mouse-ear
cross
Columbia
DATE 07-Aug-1998 #sequence-revision 03-Aug-1998 #ext-change
05-Nov-1998
D71419
ACCESSIONS
A71400
REFERENCE
#Title
Barkan, M., Barkan, P., Bock, E., Decker, F., Hoffman, H.,
Dean, C., Bergkamp, P., Dirksen, W., Van Staveren, M.,
Srikuma, K., Frost, J., Sidley, E., Hudson, S. A., Patel,
K. M., Rhyu, S., Piffanelli, P., Wedler, H., Weller, E.,
Wambutt, R., Weizenecker, T., Pohl, T. M., Terry, N.,
Gleien, J., Villarroel, R., De Clerck, P., Van Montagu, M.,
Kesseler, A., Albert, S., Gys, I., Kreis, M., Tao, N.,
Kavanagh, T., Hengge, S., Walter, P., Fulton, R., Plegier,
M., Schaeffer, M., Funk, B., Mueller-Auer, S., Silvey, M.,
James, R., Montfort, A., Rous, A., Faddameedeh, P., Douka,
A., Vouklatou, E., Milioni, D., Hatzopoulos, P.,
Piravandi, E., Obermaier, B., Hilbert, H., Duesterhoft, A.,
Moore, T., Jones, J. D. G., Eneva, T., Palme, K., Benes, V.,
Pechman, S., Ansoez, W., Cooke, R., Berber, C., Helsen,
M., Voet, M., Volckaert, G., Mewes, H. W., Kisterman, S.,
Schueller, C., Chalvatzis, N.
#Journal
Nature (1998) 391:485-488
#Title
Analysis of 1.9 Mb of contiguous sequences from Arabidopsis 4
of Arabidopsis thaliana
#cross-references MIM:6012113
#Accession
D71419
#status
preliminary, nucleic acid sequence not shown:
translation not shown
#molecule_type
DNA
#residues
1-484 #label
BEV
#cross-references MIM:6012113
GENETICS
#ref-position 4009-403845
SUMMARY
length 440 #molecular-weight 53876 #checksum 3179
Query Match 95.79 Score 45 PFC length 440:
Best Local Similarity 83.88 PFC No 5,480-00:
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 316 VWRPP 321
|||||
QV 1 VWRPP 6

Search completed. Fri Feb 25 13:24:59 2000
Job time : 52 secs.

```


Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

DB 35 VIRPPP 40
 QY 1 VVRPPP 6
 Search time: 3.1 Feb 25 14:07:10 2000
 Job time: 45 secs

```

RESULT 3
ID G3PL_MOUSE STANDARD: PRT: 440 AA.
AC Q6445: 160600:
DE CC-REV: 1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DI 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, TESTIS-SPECIFIC
DE (EC 1.2.1.12) (GAPDH).
GN GAPDS OR GAPD-S.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=TESTIS;
RX MEDLINE: 92023722
RA WELCH J.E., SCHATTE E.C., O'BRIEN D.A., EDDY E.M.:
RT "Expression of a glyceraldehyde 3-phosphate dehydrogenase gene
  specific to mouse spermatogenic cells.";
RL RIGL. REPROD. 45:869-878(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR SWISS; TISSUE=TESTIS;
RX MEDLINE: 95254745.
RA WELCH J.E., BROWN P.R., O'BRIEN D.A., EDDY E.M.:
RT "Genomic organization of a mouse glyceraldehyde 3-phosphate
  dehydrogenase gene (gapd-s) expressed in post-meiotic spermatogenic
  cells.";
RL DEV. GENET. 16:179-189(1995).
CC -1- FUNCTION MAY PLAY AN IMPORTANT ROLE IN REGULATING THE SWITCH
  BETWEEN DIFFERENT PATHWAYS FOR ENERGY PRODUCTION DURING
  SPERMIOGENESIS AND IN THE SPERMATOZOON.
CC -1- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + H2O + PHOSPHATE
  -> D-GLYCERATE + NADH.
CC -1- NAD(+) = 1,3-DIPHOSPHATEGLYCERATE + NADH.
CC -1- PATHWAY: FIRST STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: TESTIS-SPECIFIC.
CC -1- DEVELOPMENTAL STAGE: FIRST EXPRESSED AT DAY 20 IN POST-MEIOIC
  GERM CELLS. LEVELS INCREASE UNTIL DAY 24 AND THEN REMAIN CONSTANT
  DURING MATURITY.
CC -1- This Swiss PDB entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed, changed, or for commercial
  entities requires a license agreement. (See: http://www.isdb.bcm.tmc.edu/)
  or send an email to: license@isdb.bcm.tmc.edu)
CC -----
CC ENRL: M60978; G567204;
CC EMBL: U09964; G497413;
CC MGD: MGI:95653; GAPDS.
CC PROSITE: PS00071; GAPDH; 1.
CC PFAM: PF00044; gapdh; 1.
CC KW GLYCOLYSIS; OXIDOREDUCTASE; NAD.
CC FT DOMAIN 21 41 CYS/PRO-RICH.
CC FT DOMAIN 54 73 POLY-PRO.
CC FT DOMAIN 84 100 POLY-PRO.
CC FT BINDING 256 256 GLYCERALDEHYDE 3-PHOSPHATE (BY
  SIMILARITY).
CC FT ACT_SITE 263 283 ACTIVATES THIOL GROUP DURING CATALYSIS
  (BY SIMILARITY).
CC FT CONFLICT 33 34 MISSING (IN REF. 2).
CC FT CONFLICT 43 43 L -> V (IN REF. 2).
CC SQ SEQUENCE 440 AA; 47657 MW; 40693A25 CPC32;
  Query Match 95.7%; Score 45; DB 1; Length 440;
  Best Local Similarity 83.3%; Pred. No. 1.42e-00;

```


Matches 5 Conservative 1 Mismatches 0 Indels 0 Gaps 0;

DB 52 IVRPPP 67
111111
QY 1 VVRPPP 6

RESULT 4
ID 018512 PRELIMINARY: PRI: 445 AA.
AC 018512:
DI 01-JAN-1998 (PRELIMEL 05, CREATED)
DT 01-JAN-1998 (PRELIMEL 05, LAST SEQUENCE UPDATE)
DI 01-JAN-1998 (PRELIMEL 05, LAST SEQUENCE UPDATE)
DI 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)
DE SEQUONIN PECTOR PRETORUS.
OS CAENOHARDIIS ELEGANS.
OC EUKARYOTA: METAZOA: NEMATODA: SOTERINENTIA: PHAROSITIA: PHAROSITIA:
OC PHAROSITIA: PHAROSITIA: PHAROSITIA: PHAROSITIA: PHAROSITIA: PHAROSITIA:
RN [1]
RF SEQUENCE FROM N.A.
RA OLDE B., MCCOMBIE B.W.:
RL J. MOL. NEUROSCI. 9:6-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA DUBNICK M.:
F: SEQUENCING (1997) 1. 1998. JENKINS, J.P., CAIA BANKS
C: SEQUENCING (1997) 1. 1998. JENKINS, J.P., CAIA BANKS
OR ENBL. 015167. 02317845.
DR PROSITE: PS00375. 710-11.
DR PFAM: PF00001. 710-11.
FW 2. JEFFREY, J. 1997. TRANSMEMBRANE. N.Y. 1997.
SQ SEQUENCE 445 AA. 50141 MW. 458789 CPO32.

Query Match 95.7% Score 45; DB 5; Length 445;
Best Local Similarity 83.3%. Prod. No. 2.17e-00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 253 IVRPPP 208
111111
QY 1 VVRPPP 6

RESULT 5
ID 023400 PRELIMINARY: PRI: 484 AA.
AC 023400:
DI 01-JAN-1998 (PRELIMEL 05, CREATED)
DT 01-JAN-1998 (PRELIMEL 05, LAST SEQUENCE UPDATE)
DI 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)
DE INDOL-3-ACETATE PETA-GLUCOSYLTRANSFERASE.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA: VIRIDIPHYTES: STREPTOPHYTES: EUDICOTYLEDONS: ROSIDAE:
OC EUPHYLLOPHYTES: SPECIATOPHYTES: MACROPHYTES: EUDICOTYLEDONS: ROSIDAE:
OC CAPPARALES: BRASSICACEAE: ARABIDOPSIS.
RN [1]
RF SEQUENCE FROM N.A.
RA REYAN M., STUTKEMA W., MURPHY G., WARETT R., POHL T., TERRY N.,
RA KEIS M., KAVANAGH T., ENIAN K.D., RIEGER M., JAMES R.,
RA FUGIO-MENEGH P., HAIZOPOULOS P., OBERMAIER B., DIESERKOF A.,
RA JONES J., PALME K., ANSOFER W., PRISNY M., RANCIOT I., MEWIS H.W.,
RA SCHUELLER C., CHALWATZIS N.,
F: SEQUENCING (1997) 1. 1998. JENKINS, J.P., CAIA BANKS.
RN [2]
RF SEQUENCE FROM N.A.
RA EU ARABIDOPSIS SEQUENCING PROJECT, ESSA:
F: SEQUENCING (1997) 1. 1998. JENKINS, J.P., CAIA BANKS.
OR ENBL. 023400. 0302472.
DR PROSITE: PS00375. 710-11.
DR PFAM: PF00001. 710-11.
FW 2. JEFFREY, J. 1997. TRANSMEMBRANE. N.Y. 1997.
SQ SEQUENCE 484 AA. 50875 MW. 580189 CPO32.

Query Match 95.7% Score 45; DB 10; Length 484;
Best Local Similarity 83.3%. Prod. No. 2.17e-00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 315 IVRPPP 321
111111
QY 1 VVRPPP 6

RESULT 6
ID 013024 PRELIMINARY: PRI: 873 AA.
AC 013024:
DI 01-JUL-1997 (PRELIMEL 04, CRAFTED)
DT 01-JUL-1997 (PRELIMEL 04, LAST SEQUENCE UPDATE)
DI 01-JUL-1997 (PRELIMEL 04, LAST SEQUENCE UPDATE)
DI 01-NOV-1998 (PRELIMEL 08, LAST ANNOTATION UPDATE)
DE XL-INCENP.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: AMPHIBIA: BATRACHIA: ANURA:
OC MESOBRACHIA: PIPIDAE: PIPIDAE: XENOPUS.
RN [1]
RF SEQUENCE FROM N.A.
RA MEDLINE: 9727270.
RA STRENNER P.T., DUSTIN K.D., MCCARTY T.J., KING R.W., KUANG J.,
RA KIRSCHNER M.W.:
F: "Systematic identification of mitotic phosphoproteins."
PL CURR. BIOL. 7:338-348(1997).
OR ENBL. 013024. 0302472.
F: SEQUENCING (1997) 1. 1998. JENKINS, J.P., CAIA BANKS.

Query Match 95.7% Score 45; DB 13; Length 873;
Best Local Similarity 83.3%. Prod. No. 2.17e-00;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 445 IVRPPP 451
111111
QY 1 VVRPPP 6

Search Completed: Fri Feb 25 13:29:14 2000
Job time : 76 secs.